(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2022-0167288
(43) 공개일자 2022년12월20일

(51) 국제특허분류(Int. Cl.)
C07K 14/47 (2006.01) *A61K 39/00* (2006.01)
A61P 35/00 (2006.01) *C07K 14/725* (2006.01)
C07K 14/74 (2006.01)
(52) CPC특허분류
C07K 14/4748 (2013.01)
A61K 39/0011 (2021.08)
(21) 출원번호 10-2022-7036244
(22) 출원일자(국제) 2021년03월15일
심사청구일자 없음
(85) 번역문제출일자 2022년10월18일
(86) 국제출원번호 PCT/CA2021/050340
(87) 국제공개번호 WO 2021/207823
국제공개일자 2021년10월21일
(30) 우선권주장
63/009,853 2020년04월14일 미국(US)

(71) 출원인
유니버시떼 드 몬트리얼
캐나다 퀘백 에이치3티 1제이4, 몬트리얼, 에두아르드-몽쁘띠 2900
(72) 발명자
페라울트, 클라우드
캐나다, 퀘벡 에이치2브이 2엑스1, 몬트리올 애비뉴 클라우드-샴페인, 70
티볼트, 피에르
캐나다, 퀘백 에이치2엠 2엠8, 몬트리올 루에 조셉 퀸탈, 8560
(뒷면에 계속)
(74) 대리인
특허법인이룸리온

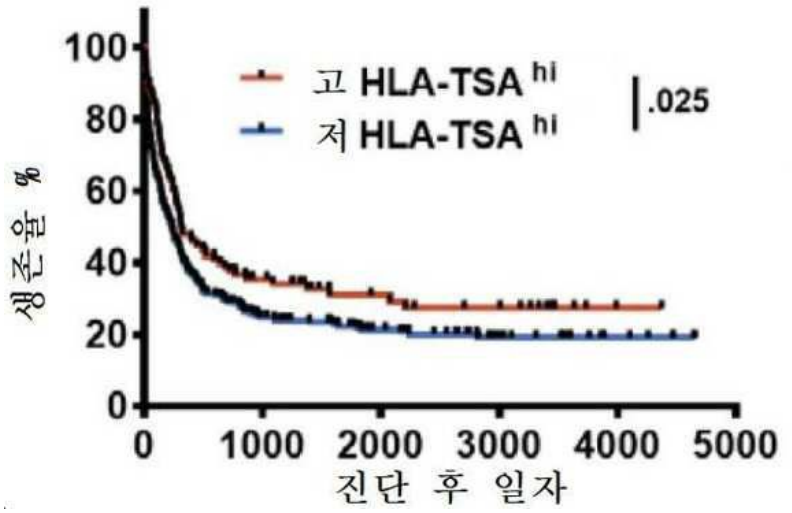
전체 청구항 수 : 총 67 항

(54) 발명의 명칭 **급성 골수성 백혈병(AML)에 대한 신규한 종양-특이적 항원 및 이의 용도**

(57) 요 약

급성 골수성 백혈병(acute myeloid leukemia, AML)은, 주로 실행 가능한 면역 표적의 결여 때문에 획기적인 면역요법으로 이익을 얻지 못하였다. 많은 비율의 AML 세포에 의해 공유되는 신규한 종양-특이적 항원(tumor-specific antigen, TSA)이 본 명세서에 기재되어 있다. 본 명세서에 기재된 TSA의 대부분은 비정상적으로 발현된 돌연변이되지 않은 게놈 서열, 예컨대, 정상 조직에서 발현되지 않는 인트론 및 유전자간 서열로부터 유래된다. 이러한 TSA로부터 유래된 핵산, 조성물, 세포 및 백신이 기재되어 있다. AML과 같은 백혈병의 치료를 위한 TSA, 핵산, 조성물, 세포 및 백신의 용도가 또한 기재되어 있다.

대 표 도 - 도5a



(52) CPC특허분류
A61P 35/00 (2018.01)
C07K 14/7051 (2013.01)
C07K 14/70539 (2013.01)
(72) 발명자
레미육스, 세바스티앵
캐나다, 퀘백 에이치8알 4씨6, 루에 뒤 포트-레미 라잘르, 271
엑스, 그레고리
벨기에, 리에주 4000 비티이 32 블러바드 드 라 소베니에르 60
하디, 마리-피에르
캐나다, 퀘백 제이6와이 0에이5, 떼흐본느 루에 드 베투시, 618

명 세 서

청구범위

청구항 1

하기 아미노산 서열 중 하나를 포함하는, 백혈병 종양 항원 펩타이드(tumor antigen peptide, TAP):

서열 (서열번호)	서열 (서열번호)	서열 (서열번호)	서열 (서열번호)
RQISVQASL (1)	SIQRNLSL (49)	AQDIILQAV (97)	RYLANKIHI (145)
DRELRNLEL (2)	NVSSHVHTV (50)	PPRPLGAQV (98)	SLLSGLLRA (146)
GARQQIHSW (3)	ALASHLIEA (51)	FNVALNARY (99)	SRIHLVVL (147)
SGKLRVAL (4)	ALDDITIQL (52)	GPGSRESTL (100)	SSSPVRGPSV (148)
RSASSATQVHK (5)	ALGNTVPAV (53)	IPHQRSSL (101)	STFSLYLKK (149)
SASSATQVHK (6)	ALLPAVPSL (54)	LTDRIYLTL (102)	SLDLLPLSI (150)
FLLEFKPVS (7)	APAPPPVAV (55)	NLKEKKALF (103)	VTDLLALTV (151)
GPQVRGSI (8)	APDKKITL (56)	VLFGGKVSGA (104)	RTQITKVSLKK (152)
IRMKAQAL (9)	AQMNLLQKY (57)	VVFPFPVNK (105)	ILRSPLKW (153)
KIKVFSKVY (10)	DQVIRLAGL (58)	SLLIIPKKK (106)	LSTGHLSTV (154)
LLSRGLLFRI (11)	ETTSQVRKY (59)	APGAAGQRL (107)	TVEEYLVNI (155)
LPIASASLL (12)	GGSLIHPQW (60)	KLQDKEIGL (108)	QIKTKLLGSL (156)
LYFLGHGSI (13)	GLYYKLHNV (61)	SLREPQPAL (109)	LPSFSHFLLL (157)
NPLQLSLSI (14)	GQKPVILTY (62)	TPGRSTQAI (110)	CLRIGPVTL (158)
DLMLRESL (15)	GSLDFQRGW (63)	APRGTAAL (111)	HVSDGSTALK (159)
VTFKLSLF (16)	HHLVETLKF (64)	IASPIALL (112)	IAYSVRALR (160)
IALYKQVL (17)	HLLSETPQL (65)	ILFQNSALK (113)	PRGFLSAL (161)
IVATGSLLK (18)	HQLYRASAL (66)	ILKKNISI (114)	ISSWLISSL (162)
KIKNKTKNK (19)	HTDDIENAKY (67)	IPLAVRTI (115)	IPLNPFSSL (163)
KLLSLTIYK (20)	IAAPILHV (68)	LPRNKPLL (116)	LSDRQLSL (164)
NILKKTVL (21)	KAFPFHIIF (69)	PAPPHPAAL (117)	LSHPAPSSL (165)
NPKLKDIL (22)	KATEYVHSL (70)	SPVVRVGL (118)	LRKAVDPIL (166)
NQKKVRIL (23)	KFSNVTMLF (71)	TLNQGINVYI (119)	ILLEEQSLI (167)
PFPLVQVEPV (24)	KLLEKAFSI (72)	RPRGPRTAP (120)	LTSISIRPV (168)
SPQSGPAL (25)	KPMPTKVVF (73)	SVQLLEQAIHK (121)	TISECPLLI (169)
TSRLPKIQK (26)	NVNRPLTMK (74)	RTPKNYQHW (122)	TLKLKKIFF (170)
LLDNILQSI (27)	REPYELTVPAL (75)	ALPVALPSL (123)	ILLSNFSSL (171)
RLEVRKVIL (28)	SEAEAAKNAL (76)	SLQILVSSL (124)	LGGAWKAVF (172)
LSWGYFLFK (29)	SLWGQPAEA (77)	ISNKVPKLF (125)	LSASHLSSL (173)
TILPRILTL (30)	SPADHRGYASL (78)	TVIRIAIVNK (126)	AGDIIARLI (174)
EGKIKRNI (31)	SPQSAAAEL (79)	KEIFLELRL (127)	DRGILRNLL (175)
FLASFVEKTVL (32)	SPVVHQSL (80)	TLRSPGSSL (128)	GLRLIHVSL (176)
ILASHNLTV (33)	SPYRTPVL (81)	TVRGDVSSL (129)	GLRLLHVSL (177)
IQLTSVHLL (34)	SVFAGVVGV (82)	ALDPLLLRI (130)	LHNEKGLSL (178)
LELISFLPVL (35)	SYSPAHARL (83)	ISLIVTGLK (131)	LPSFSRPSGII (179)
NFCMLHQSI (36)	THGSEQLHL (84)	KILDVNLRI (132)	LSSRLPLGK (180)
PARPAGPL (37)	TQAPPNVVL (85)	ERVYIRASL (133)	MIGIKRLL (181)
PLPIVPAL (38)	VLVPYEPPQV (86)	ILDLESRY (134)	NLKKREIL (182)
SNLIRTGSH (39)	VSFPDVRKV (87)	KTFVQQKTL (135)	RMVAYLQQL (183)
VPAPAQAI (40)	VVFDKSDLAKY (88)	LYIKSLPAL (136)	SPARALPSL (184)
KGHGGPRSW (41)	YSHHSGLEY (89)	VLKEKNASL (137)	TVPGIQRY (185)
ITSSAVTTALK (42)	YYLDWIHHY (90)	LGISLTLKY (138)	VSRNYVLLI (186)
LLLPESPSI (43)	SVYKYLKAK (91)	DLLPKKLL (139)	LTVPLSVFW (187)
VILIPLPPK (44)	IYQFIMDRF (92)	HSLISIVYL (140)	KLNQAFLVL (188)
AVLLPKPPK (45)	GTLQGIRAW (93)	IAGALRSVL (141)	RLVSSTLLQK (189)
TQVSMAESI (46)	AQKVSVGQAA (94)	IGNPILRVL (142)	LPSHSLLI (190)
LNHLRTSI (47)	LYPSKLTHF (95)	IYAPHIRLS (143)	
NTSHLPLIY (48)	ATQNTIIGK (96)	LRSQILSY (144)	

청구항 2

제1항에 있어서, 서열번호 97 내지 154에 제시된 아미노산 서열 중 하나를 포함하는, 백혈병 TAP.

청구항 3

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*01:01 분자에 결합하고 그리고 아미노산 서열 NTSHLPLIY(서열번호 48), HTDDIENAKY(서열번호 67), YSHHSGLEY(서열번호 89), ILDLESRY(서열번호 134), VTDLLALTV(서열번호 151) 또는 LSDRQLSL(서열번호 164), 바람직하게는 ILDLESRY(서열번호 134) 또는 VTDLLALTV(서열번호 151)를 포함하는, 백혈병 TAP.

청구항 4

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*02:01 분자에 결합하고 그리고 아미노산 서열 FLLEFKPVS(서열번호 7), LLSRGLLFRI(서열번호 11), LLDNILQSI(서열번호 27), FLASFVEKTVL(서열번호 32), ILASHNLTV(서열번호 33), IQLTSVHLL(서열번호 34), LELISFLPVL(서열번호 35), LLLPESPSI(서열번호 43), ALASHLIEA(서열번호 51), ALDDITIQL(서열번호 52), ALGNTVPAV(서열번호 53), ALLPAVPSL(서열번호 54), GLYYKLHNV(서열번호 61), HLLSETPQL(서열번호 65), KLLEKAFSI(서열번호 72), SLWGQPAEA(서열번호 77), SVFAGVVGV(서열번호 82), VLVPYEPPQV(서열번호 86), VLFGGKVSGA(서열번호 104), KLQDKEIGL(서열번호 108), TLNQGINVYI(서열번호 119), ALPVALPSL(서열번호 123), ALDPLLLRI(서열번호 130), KILDVNLRI(서열번호 132), SLLSGLLRA(서열번호 146), SLDLLPLSI(서열번호 150), ILLEEQSLI(서열번호 167), LTSISIRPV(서열번호 168), TISECPLLI(서열번호 169), ILLSNFSSL(서열번호 171), RMVAYLQQL(서열번호 183), 또는 KLNQAFLVL(서열번호 188), 바람직하게는 VLFGGKVSGA(서열번호 104), KLQDKEIGL(서열번호 108), TLNQGINVYI(서열번호 119), ALPVALPSL(서열번호 123), ALDPLLLRI(서열번호 130), KILDVNLRI(서열번호 132), SLLSGLLRA(서열번호 146) 또는 SLDLLPLSI(서열번호 150)를 포함하는, 백혈병 TAP.

청구항 5

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*03:01 분자에 결합하고 그리고 아미노산 서열 RSASSATQVHK(서열번호 5), IVATGSLLK(서열번호 18), KIKNKTKNK(서열번호 19), KLLSLTIYK(서열번호 20), ITSSAVTTALK(서열번호 42), VILIPLPPK(서열번호 44), NVNRPLTMK(서열번호 74), SVYKYLKAK(서열번호 91), VVFPFPVNK(서열번호 105), ILFQNSALK(서열번호 113), TVIRIAIVNK(서열번호 126), ISLIVTGLK(서열번호 131), HVSDGSTALK(서열번호 159), IAYSVRALR(서열번호 160), LSSRLPLGK(서열번호 180) 또는 RLVSSTLLQK(서열번호 189), 바람직하게는 VVFPFPVNK(서열번호 105), ILFQNSALK(서열번호 113), TVIRIAIVNK(서열번호 126) 또는 ISLIVTGLK(서열번호 131)를 포함하는, 백혈병 TAP.

청구항 6

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*11:01 분자에 결합하고 그리고 아미노산 서열 SASSATQVHK(서열번호 6), AVLLPKPPK(서열번호 45), ATQNTIIGK(서열번호 96), SLLIIPKKK(서열번호 106), SVQLLEQAIHK(서열번호 121), STFSLYLKK(서열번호 149) 또는 RTQITKVSLKK(서열번호 152), 바람직하게는 SLLIIPKKK(서열번호 106), SVQLLEQAIHK(서열번호 121), STFSLYLKK(서열번호 149) 또는 RTQITKVSLKK(서열번호 152)를 포함하는, 백혈병 TAP.

청구항 7

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*24:02 분자에 결합하고 그리고 아미노산 서열 LYFLGHGSI(서열번호 13), NFCMLHQSI(서열번호 36), KFSNVTMLF(서열번호 71), IYQFIMDRF(서열번호 92), LYPSKLTHF(서열번호 95) 또는 RYLANKIHI(서열번호 145), 바람직하게는 RYLANKIHI(서열번호 145)를 포함하는, 백혈병 TAP.

청구항 8

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*26:01 분자에 결합하고 그리고 아미노산 서열 ETTSQVRKY(서열번호 59) 또는 TVPGIQRY(서열번호 185)를 포함하는, 백혈병 TAP.

청구항 9

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*29:02 분자에 결합하고 그리고 아미노산 서열 VVFDKSDLAKY(서열번호 88), FNVALNARY(서열번호 99) 또는 LGISLTLKY(서열번호 138), 바람직하게는 FNVALNARY(서열번호 99) 또는 LGISLTLKY(서열번호 138) 중 하나를 포함하는, 백혈병 TAP.

청구항 10

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*30:01 분자에 결합하고 그리고 아미노산 서열 TSRLPKIQK(서열번호 26), LSWGYFLFK(서열번호 29) 또는 LSHPAPSSL(서열번호 165)을 포함하는, 백혈병 TAP.

청구항 11

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-A*68:02 분자에 결합하고 그리고 아미노산 서열 NVSSHVHTV(서열번호 50) 또는 SSSPVRGPSV(서열번호 148), 바람직하게는 SSSPVRGPSV(서열번호 148)를 포함하는, 백혈병 TAP.

청구항 12

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*07:02 분자에 결합하고 그리고 아미노산 서열 GPQVRGSI(서열번호 8), SPQSGPAL(서열번호 25), VPAPAQAI(서열번호 40), APAPPPVAV(서열번호 55), APDKKITL(서열번호 56), KPMPTKVVF(서열번호 73), SPADHRGYASL(서열번호 78), SPQSAAAEL(서열번호 79), SPVVHQSL(서열번호 80), SPYRTPVL(서열번호 81), PPRPLGAQV(서열번호 98), GPGSRESTL(서열번호 100), APGAAGQRL(서열번호 107), TPGRSTQAI(서열번호 110), APRGTAAL(서열번호 111), SPVVRVGL(서열번호 118), RPRGPRTAP(서열번호 120), TLRSPGSSL(서열번호 128), TVRGDVSSL(서열번호 129), LPSFSHFLLL(서열번호 157), PRGFLSAL(서열번호 161), IPLNPFSSL(서열번호 163), LPSFSRPSGII(서열번호 179) 또는 SPARALPSL(서열번호 184), 바람직하게는 PPRPLGAQV(서열번호 98), GPGSRESTL(서열번호 100), APGAAGQRL(서열번호 107), TPGRSTQAI(서열번호 110), APRGTAAL(서열번호 111), SPVVRVGL(서열번호 118), RPRGPRTAP(서열번호 120), TLRSPGSSL(서열번호 128) 또는 TVRGDVSSL(서열번호 129)을 포함하는, 백혈병 TAP.

청구항 13

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*08:01 분자에 결합하고 그리고 아미노산 서열 SGKLRVAL(서열번호 4), NPLQLSLSI(서열번호 14), DLMLRESL(서열번호 15), IALYKQVL(서열번호 17), NILKKTVL(서열번호 21), NPKLKDIL(서열번호 22), NQKKVRIL(서열번호 23), RLEVRKVIL(서열번호 28), EGKIKRNI(서열번호 31), LNHLRTSI(서열번호 47), SIQRNLSL(서열번호 49), IPHQRSSL(서열번호 101), NLKEKKALF(서열번호 103), ILKKNISI(서열번호 114), VLKEKNASL(서열번호 137), DLLPKKLL(서열번호 139), SRIHLVVL(서열번호 147), QIKTKLLGSL(서열번호 156), TLKLKKIFF(서열번호 170), MIGIKRLL(서열번호 181) 또는 NLKKREIL(서열번호 182), 바람직하게는 IPHQRSSL(서열번호 101), NLKEKKALF(서열번호 103), ILKKNISI(서열번호 114), VLKEKNASL(서열번호 137), DLLPKKLL(서열번호 139) 또는 SRIHLVVL(서열번호 147)을 포함하는, 백혈병 TAP.

청구항 14

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*14:01 분자에 결합하고 그리고 아미노산 서열 DRELRNLEL(서열번호 2), SNLIRTGSH(서열번호 39), DQVIRLAGL(서열번호 58), HQLYRASAL(서열번호 66), SLQILVSSL(서열번호 124), ERVYIRASL(서열번호 133), LYIKSLPAL(서열번호 136), IAGALRSVL(서열번호 141), ISSWLISSL(서열번호 162), DRGILRNLL(서열번호 175), GLRLIHVSL(서열번호 176) 또는 GLRLLHVSL(서열번호 177), 바람직하게는 SLQILVSSL(서열번호 124), ERVYIRASL(서열번호 133), LYIKSLPAL(서열번호 136) 또는 IAGALRSVL(서열번호 141)을 포함하는, 백혈병 TAP.

청구항 15

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*15:01 분자에 결합하고 그리고 아미노산 서열 KIKVFSKVY(서열번호 10), AQMNLLQKY(서열번호 57), GQKPVILTY(서열번호 62) 또는 AQKVSVGQAA(서열번호 94)를 포함하는, 백혈병 TAP.

청구항 16

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*27:05 분자에 결합하고 그리고 아미노산 서열 RQISVQASL(서열번호 1) 또는 LRSQILSY(서열번호 144), 바람직하게는 LRSQILSY(서열번호 144)를 포함하는, 백혈병 TAP.

청구항 17

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*38:01 분자에 결합하고 그리고 아미노산 서열 TQVSMAESI(서열번호 46), HHLVETLKF(서열번호 64) 또는 THGSEQLHL(서열번호 84)을 포함하는, 백혈병 TAP.

청구항 18

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*40:01 분자에 결합하고 그리고 아미노산 서열 REPYELTVPAL(서열번호 75) 또는 SEAEAAKNAL(서열번호 76)을 포함하는, 백혈병 TAP.

청구항 19

제1항 또는 제2항에 있어서, 상기 백혈병 TAP는 HLA-B*44:03 분자에 결합하고 그리고 아미노산 서열 KEIFLELRL(서열번호 127)을 포함하는, 백혈병 TAP.

청구항 20

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*51:01 분자에 결합하고 그리고 아미노산 서열 LPIASASLL(서열번호 12), PFPLVQVEPV(서열번호 24), PLPIVPAL(서열번호 38), IAAPILHV(서열번호 68), IPLAVRTI(서열번호 115), LPRNKPLL(서열번호 116) 또는 LPSHSLLI(서열번호 190), 바람직하게는 IPLAVRTI(서열번호 115) 또는 LPRNKPLL(서열번호 116)을 포함하는, 백혈병 TAP.

청구항 21

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*57:01 분자에 결합하고 그리고 아미노산 서열 GARQQIHSW(서열번호 3), VTFKLSLF(서열번호 16), KGHGGPRSW(서열번호 41), GSLDFQRGW(서열번호 63), KAFPFHIIF(서열번호 69), GTLQGIRAW(서열번호 93), RTPKNYQHW(서열번호 122), ISNKVPKLF(서열번호 125), KTFVQQKTL(서열번호 135), ILRSPLKW(서열번호 153) 또는 LTVPLSVFW(서열번호 183), 바람직하게는 RTPKNYQHW(서열번호 122), ISNKVPKLF(서열번호 125), KTFVQQKTL(서열번호 135) 또는 ILRSPLKW(서열번호 153)를 포함하는, 백혈병 TAP.

청구항 22

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-B*57:03 분자에 결합하고 그리고 아미노산 서열 GGSLIHPQW(서열번호 60) 또는 LGGAWKAVF(서열번호 172)를 포함하는, 백혈병 TAP.

청구항 23

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*03:03 분자에 결합하고 그리고 아미노산 서열 PARPAGPL(서열번호 37), IASPIALL(서열번호 112) 또는 HSLISIVYL(서열번호 140), 바람직하게는 IASPIALL(서열번호 112) 또는 HSLISIVYL(서열번호 140)을 포함하는, 백혈병 TAP.

청구항 24

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*05:01 분자에 결합하고 그리고 아미노산 서열 SLDLLPLSI(서열번호 150)를 포함하는, 백혈병 TAP.

청구항 25

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*06:02 분자에 결합하고 그리고 아미노산 서열 IRMKAQAL(서열번호 9), KATEYVHSL(서열번호 70), VSFPDVRKV(서열번호 87), IGNPILRVL(서열번호 142), LSTGHLSTV(서열번호 154) 또는 LRKAVDPIL(서열번호 166), 바람직하게는 IGNPILRVL(서열번호 142) 또는 LSTGHLSTV(서열번호 154)를 포함하는, 백혈병 TAP.

청구항 26

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*07:01 분자에 결합하고 그리고 아미노산 서열 IGNPILRVL(서열번호 142), IYAPHIRLS(서열번호 143), TVEEYLVNI(서열번호 155), LHNEKGLSL(서열번호 178) 또는 VSRNYVLLI(서열번호 186), 바람직하게는 IGNPILRVL(서열번호 142) 또는 IYAPHIRLS(서열번호 143)를 포함하는, 백혈병 TAP.

청구항 27

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*07:02 분자에 결합하고 그리고 아미노산 서열 TILPRILTL(서열번호 30), SYSPAHARL(서열번호 83), TQAPPNVVL(서열번호 85), YYLDWIHHY(서열번호 90), SLREPQPAL(서열번호 109), PAPPHPAAL(서열번호 117) 또는 CLRIGPVTL(서열번호 158), 바람직하게는 SLREPQPAL(서열번호 109) 또는 PAPPHPAAL(서열번호 117)을 포함하는, 백혈병 TAP.

청구항 28

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*08:02 분자에 결합하고 그리고 아미노산 서열 AQDIILQAV(서열번호 97), LTDRIYLTL(서열번호 102) 또는 AGDIIARLI(서열번호 174), 바람직하게는 AQDIILQAV(서열번호 97) 또는 LTDRIYLTL(서열번호 102)을 포함하는, 백혈병 TAP.

청구항 29

제1항 또는 제2항에 있어서,

상기 백혈병 TAP는 HLA-C*12:03 분자에 결합하고 그리고 아미노산 서열 LSASHLSSL(서열번호 173)을 포함하는, 백혈병 TAP.

청구항 30

제1항 내지 제29항 중 어느 한 항에 있어서,

게놈의 비-단백질 코딩 영역에 존재하는 서열에 의해 인코딩되는, 백혈병 TAP.

청구항 31

제30항에 있어서,

상기 게놈의 상기 비-단백질 코딩 영역은 비번역 전사된 영역(untranslated transcribed region, UTR)인, 백혈병 TAP.

청구항 32

제30항에 있어서,

상기 게놈의 상기 비-단백질 코딩 영역은 인트론인, 백혈병 TAP.

청구항 33

제30항에 있어서,

상기 게놈의 상기 비-단백질 코딩 영역은 유전자간 영역인, 백혈병 TAP.

청구항 34

제1항 내지 제33항 중 어느 한 항에 정의된 백혈병 TAP의 적어도 2가지를 포함하는, 조합물.

청구항 35

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP 또는 제34항의 조합물을 인코딩하는, 핵산.

청구항 36

제35항에 있어서,

mRNA 또는 바이러스 벡터인, 핵산.

청구항 37

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP, 제34항의 조합물, 또는 제35항 또는 제36항의 핵산을 포함하는, 리포솜.

청구항 38

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP, 제34항의 조합물, 제35항 또는 제36항의 핵산, 또는 제37항의 리포솜, 및 약제학적으로 허용 가능한 담체를 포함하는, 조성물.

청구항 39

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP, 제34항의 조합물, 제35항 또는 제36항의 핵산, 제37항의 리포솜, 또는 제38항의 조성물 및 애주번트(adjuvant)를 포함하는, 백신.

청구항 40

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP를 펩타이드 결합 홈에 포함하는, 단리된 주조직적합성 복합체(major histocompatibility complex, MHC) 클래스 I 분자.

청구항 41

제40항에 있어서,

다량체의 형태인, 단리된 MHC 클래스 I 분자.

청구항 42

제41항에 있어서,

상기 다량체는 사량체인, 단리된 MHC 클래스 I 분자.

청구항 43

단리된 세포로서,

(i) 제1항 내지 제33항 중 어느 한 항의 백혈병 TAP, (ii) 제34항의 조합물 또는 (iii) 제1항 내지 제33항 중 어느 한 항의 TAP 또는 제34항의 조합물을 인코딩하는 뉴클레오타이드 서열을 포함하는 벡터를 포함하는, 단리된 세포.

청구항 44

단리된 세포로서,

제1항 내지 제33항 중 어느 한 항의 백혈병 TAP 또는 제34항의 조합물을 펩타이드 결합 홈에 포함하는 주조직적 합성 복합체(major histocompatibility complex, MHC) 클래스 I 분자를 표면에서 발현하는, 단리된 세포.

청구항 45

제44항에 있어서,

항원-제시 세포(antigen-presenting cell, APC)인, 세포.

청구항 46

제45항에 있어서,

상기 APC는 수지상 세포인, 세포.

청구항 47

제40항 내지 제42항 중 어느 한 항의 단리된 MHC 클래스 I 분자 및/또는 제44항 내지 제46항 중 어느 한 항의 세포의 표면에서 발현되는 MHC 클래스 I 분자를 특이적으로 인식하는, T-세포 수용체(T-cell receptor, TCR).

청구항 48

제47항에 있어서,

상기 TCR은 서열번호 191 내지 219에 제시된 아미노산 서열에 제시된 아미노산 서열 중 하나를 포함하는 상보성 결정 영역 3(CDR3)을 포함하는 TCR베타(TCRβ) 사슬을 포함하는, TCR.

청구항 49

제47항 또는 제48항의 TCR을 세포 표면에 발현하는, 단리된 세포.

청구항 50

제49항에 있어서,

$CD8^+$ T 림프구인, 단리된 세포.

청구항 51

적어도 0.5%의 제49항 또는 제50항에 정의된 바와 같은 단리된 세포를 포함하는, 세포 집단.

청구항 52

대상체에서 백혈병을 치료하는 방법으로서,

상기 대상체에게 유효량의, (i) 제1항 내지 제33항 중 어느 한 항의 백혈병 TAP; (ii) 제34항의 조합물; (iii) 제35항 또는 제36항의 핵산; (iv) 제37항의 리포솜; (v) 제38항의 조성물; (vi) 제39항의 백신; (vii) 제43항 내지 제46항, 제49항 및 제50항 중 어느 한 항의 세포; 또는 (viii) 제51항의 세포 집단;을 투여하는 단계를 포함하는, 방법.

청구항 53

제52항에 있어서,

상기 백혈병은 골수성 백혈병인, 방법.

청구항 54

제53항에 있어서,

상기 골수성 백혈병은 급성 골수성 백혈병(acute myeloid leukemia, AML)인, 방법.

청구항 55

제52항 내지 제54항 중 어느 한 항에 있어서,

상기 대상체에게 적어도 하나의 추가의 항종양 제제 또는 요법을 투여하는 단계를 더 포함하는, 방법.

청구항 56

제55항에 있어서,

상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 방법.

청구항 57

대상체에서 백혈병을 치료하기 위한, (i) 제1항 내지 제33항 중 어느 한 항의 백혈병 TAP; (ii) 제34항의 조합물; (iii) 제35항 또는 제36항의 핵산; (iv) 제37항의 리포솜; (v) 제38항의 조성물; (vi) 제39항의 백신; (vii) 제43항 내지 제46항, 제49항 및 제50항 중 어느 한 항의 세포; 또는 (viii) 제51항의 세포 집단;의 용도.

청구항 58

대상체에서 백혈병을 치료하기 위한 의약의 제조를 위한, (i) 제1항 내지 제33항 중 어느 한 항의 백혈병 TAP; (ii) 제34항의 조합물; (iii) 제35항 또는 제36항의 핵산; (iv) 제37항의 리포솜; (v) 제38항의 조성물; (vi) 제39항의 백신; (vii) 제43항 내지 제46항, 제49항 및 제50항 중 어느 한 항의 세포; 또는 (viii) 제51항의 세포 집단;의 용도.

청구항 59

제57항 또는 제58항에 있어서,

상기 백혈병은 골수성 백혈병인, 용도.

청구항 60

제59항에 있어서,

상기 골수성 백혈병은 급성 골수성 백혈병(AML)인, 용도.

청구항 61

제57항 내지 제60항 중 어느 한 항에 있어서,

적어도 하나의 추가의 항종양 제제 또는 요법의 사용을 더 포함하는, 용도.

청구항 62

제61항에 있어서,

상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 방법.

청구항 63

대상체에서 백혈병의 치료에 사용하기 위한, (i) 제1항 내지 제33항 중 어느 한 항의 백혈병 TAP; (ii) 제34항의 조합물; (iii) 제35항 또는 제36항의 핵산; (iv) 제37항의 리포솜; (v) 제38항의 조성물; (vi) 제39항의 백신; (vii) 제43항 내지 제46항, 제49항 및 제50항 중 어느 한 항의 세포; 또는 (viii) 제51항의 세포 집단.

청구항 64

제63항에 있어서,

상기 백혈병은 골수성 백혈병인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

청구항 65

제64항에 있어서,

상기 골수성 백혈병은 급성 골수성 백혈병(AML)인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

청구항 66

제63항 내지 제65항 중 어느 한 항에 있어서,

적어도 하나의 추가의 항종양 제제 또는 요법과 조합하여 사용하기 위한, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

청구항 67

제66항에 있어서,

상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

발명의 설명

기 술 분 야

[0001] **관련 출원에 대한 상호 참조**

[0002] 본 출원은 2020년 4월 14일자로 출원된 미국 특허 가출원 제63/009,853호의 이익을 주장하며, 이 기초 출원은 그 전문이 본 명세서에 참조에 의해 원용된다.

[0003] **서열 목록**

[0004] 해당 없음.

[0005] **기술 분야**

[0006] 본 발명은 일반적으로 암에 관한 것으로, 보다 구체적으로는 T-세포-기반 암면역요법에 유용한 급성 골수성 백혈병에 특이적인 종양 항원에 관한 것이다.

배 경 기 술

[0007] 가장 공격적인 혈액 악성 종양인 급성 골수성 백혈병(acute myeloid leukemia, AML)은 비정상적인 후성적 패턴화, 교란된 미토콘드리아 단백질 정체 및 상대적으로 적은 수의 돌연변이를 특징으로 하는 이질적인 질환이다(Li et al., 2016; Ntziachristos et al., 2016; Ishizawa et al., 2019; Fennell et al., 2019). 특히, AML의 유전적 및 후성적 변화는 진단보다 수년 앞서 있을 수 있다(Abelson et al., 2018; Desai et al., 2018). 또한, 치유는 대량의 종양 세포뿐만 아니라 백혈병 줄기 세포도 제거해야 할 필요가 있다(Shlush et al., 2017; Boyd et al., 2018). 현재 대부분의 환자는 화학 요법 후 재발하며, 60세 미만 환자의 경우 5년 전체 생존율은 40%이고 (AML 사례의 대부분을 나타내는) 60세 이상 환자의 경우 10 내지 20%에 불과하다(Vasu et al., 2018).

[0008] 지난 몇 년에 걸쳐서, 암 면역요법에 대한 열정은 주로 두 가지 주요 돌파구에 의해 촉진되었다: i) 흑색종 및 선택된 유형의 고형 종양의 치료를 위한 면역관문 요법 및 ii) 림프성 악성 종양의 치료를 위한 키메라 항원 수용체. 그러나, AML은 주로 실행 가능한 면역 표적이 없기 때문에 이러한 혁신의 혜택을 받지 못했다. T 세포에 의해 인식되는 주조직적합성 복합체 MHC-연관 펩타이드(MHC-associated peptide, MAP)가 항암 반응의 핵심에 있다는 개념에 따라(Coulie et al., 2014), AML 세포가 CD8 T 세포에 면역원성 MAP를 제시해야 한다는 증거가 시사된다: i) AML 세포는 고밀도의 MHC 클래스 I 분자를 발현하고(Berlin et al., 2015) 그리고 ii) AML 환자의 골수는 탈진의(따라서 항원 인식의) 표현형 및 전사 특징을 갖는 CD8 T 세포를 함유한다(Knaus et al., 2018). 그러나, 보호적 면역 반응을 유도할 수 있는 AML 항원의 속성은 여전히 포착되기 어렵다.

[0009] 암 면역학자의 관심을 끈 MAP의 첫 번째 부류는 정상 세포에 비해 종양 세포에서 과발현되는 종양-연관 항원(tumor-associated antigen, TAA)이다. 자가 항원을 인식하는 고친화도 T 세포는 흉선 선택의 중추 관용 과정에 의해 제거되기 때문에, TAA는 본질적으로 저친화도 T 세포에 의해 인식된다. 따라서, TAA-기반 백신은 AML 진화에 설득력 있는 영향을 미치지 않았다. 가장 많이 연구된 AML TAA: 윌름스 종양 1(*WT1*)에서 특히 실망스러운 결

과를 얻었다(Di Stasi et al., 2015; Maslak et al., 2018; Rashidi and Walter, 2016). 중요하게도, 최근 보고서는, *WT1*-유래 펩타이드를 표적으로 하는 TCR 유전자 요법(T 세포가 선택된 항원에 대해 고친화도 TCR을 발현하도록 조작됨)이 동종 조혈 줄기세포 이식 수용체의 재발을 지속적으로 예방할 수 있었음을 나타내었다(Chapuis et al., 2019). 전반적으로, 이러한 연구는 *WT1*-유래 펩타이드가 면역원성이 나쁘고 완전한 치료적 잠재력에 도달하기 위해 조작된 T 세포로 표적화될 필요가 있음을 시사한다.

[0010] TAA와 대조적으로, 종양 특이적 항원(tumor specific antigen, TSA)은 종양 세포에 의해 단독으로 제시되는 MAP이다. 지금까지, 신생항원으로도 알려진 돌연변이 TSA(mTSA)는 최근 고형 종양에 대한 백신 검토에서 상당한 관심을 끌고 있다. 실제로, mTSA는 중추 관용을 유도하는 수질 흉선 세포(mTEC)에서 발견되지 않기 때문에 고도로 면역원성일 수 있다. 그러나, mTSA는 두 가지 주의 사항을 제시한다. 첫째로, 이들은 일반적으로 각 환자의 종양(개인 신생항원)에 고유하다. 둘째로, 이들은 처음에 예측된 것보다 덜 일반적이다(Knaus et al., 2018). AML 세포의 낮은 돌연변이 부담과 일관되게, 1차 AML 세포의 질량 분광(MS) 분석에 의해 단 하나의 mTSA만 검증되었다(van der Lee et al., 2019). NPM1 유전자에서 프레임시프트로부터 파생된 이 mTSA의 치료 가능성은 아직 평가되지 않았지만, 사용 가능한 증거에 따르면, AML 환자에서 자발적 면역 반응을 유도하지 못한다(van der Lee et al., 2019).

[0011] 이에 비추어, AML에 대한 치료적 면역 반응을 유발할 수 있는 항원을 식별하는 것이 절실히 필요하다. 이러한 항원은 백신(±면역관문저해제)으로서 또는 T-세포 수용체-기반 접근법(세포 요법, 이중특이성 생물학적 제제)의 표적으로서 사용될 수 있었다.

[0012] 본 기술내용은 다수의 문헌을 참조하며, 이러한 문헌의 내용은 그 전문이 본 명세서에 참조에 의해 원용된다.

발명의 내용

[0013] 본 개시내용은 하기 항목 1 내지 67을 제공한다:

[0014] 1. 하기 아미노산 서열 중 하나를 포함하는, 백혈병 종양 항원 펩타이드(tumor antigen peptide, TAP):

서열 (서열번호)	서열 (서열번호)	서열 (서열번호)	서열 (서열번호)
RQISVQASL (1)	SIQRNLSL (49)	AQDIILQAV (97)	RYLANKIHI (145)
DRELRNLEL (2)	NVSSHVHTV (50)	PPRPLGAQV (98)	SLLSGLLRA (146)
GARQQIHSW (3)	ALASHLIEA (51)	FNVALNARY (99)	SRIHLVVL (147)
SGKLRVAL (4)	ALDDITIQL (52)	GPGSRESTL (100)	SSSPVRGPSV (148)
RSASSATQVHK (5)	ALGNTVPAV (53)	IPHQRSSL (101)	STFSLYLKK (149)
SASSATQVHK (6)	ALLPAVPSL (54)	LTDRIYLTL (102)	SLDLLPLSI (150)
FLLEFKPVS (7)	APAPPPVAV (55)	NLKEKKALF (103)	VTDLLALTV (151)
GPQVRGSI (8)	APDKKITL (56)	VLFGGKVSGA (104)	RTQITKVSLKK (152)
IRMKAQAL (9)	AQMNLLQKY (57)	VVFPFPVNK (105)	ILRSPLKW (153)
KIKVFSKVY (10)	DQVIRLAGL (58)	SLLIIPKKK (106)	LSTGHLSTV (154)
LLSRGLLFRI (11)	ETTSQVRKY (59)	APGAAGQRL (107)	TVEEYLVNI (155)
LPIASASLL (12)	GGSLIHPQW (60)	KLQDKEIGL (108)	QIKTKLLGSL (156)
LYFLGHGSI (13)	GLYYKLHNV (61)	SLREPQPAL (109)	LPSFSHFLLL (157)
NPLQLSLSI (14)	GQKPVILTY (62)	TPGRSTQAI (110)	CLRIGPVTL (158)
DLMLRESL (15)	GSLDFQRGW (63)	APRGTAAL (111)	HVSDGSTALK (159)
VTFKLSLF (16)	HHLVETLKF (64)	IASPIALL (112)	IAYSVRALR (160)
IALYKQVL (17)	HLLSETPQL (65)	ILFQNSALK (113)	PRGFLSAL (161)
IVATGSLLK (18)	HQLYRASAL (66)	ILKKNISI (114)	ISSWLISSL (162)
KIKNKTKNK (19)	HTDDIENAKY (67)	IPLAVRTI (115)	IPLNPFSSL (163)
KLLSLTIYK (20)	IAAPILHV (68)	LPRNKPLL (116)	LSDRQLSL (164)
NILKKTVL (21)	KAFPFHIIF (69)	PAPPHPAAL (117)	LSHPAPSSL (165)
NPKLKDIL (22)	KATEYVHSL (70)	SPVVRVGL (118)	LRKAVDPIL (166)

[0015]

NQKKVRIL (23)	KFSNVTMLF (71)	TLNQGINVYI (119)	ILLEEQSLI (167)
PFPLVQVEPV (24)	KLLEKAFSI (72)	RPRGPRTAP (120)	LTSISIRPV (168)
SPQSGPAL (25)	KPMPTKVVF (73)	SVQLLEQAIHK (121)	TISECPLLI (169)
TSRLPKIQK (26)	NVNRPLTMK (74)	RTPKNYQHW (122)	TLKLKKIFF (170)
LLDNILQSI (27)	REPYELTVPAL (75)	ALPVALPSL (123)	ILLSNFSSL (171)
RLEVRKVIL (28)	SEAEAAKNAL (76)	SLQILVSSL (124)	LGGAWKAVF (172)
LSWGYFLFK (29)	SLWGQPAEA (77)	ISNKVPKLF (125)	LSASHLSSL (173)
TILPRILTL (30)	SPADHRGYASL (78)	TVIRIAIVNK (126)	AGDIIARLI (174)
EGKIKRNI (31)	SPQSAAAEL (79)	KEIFLELRL (127)	DRGILRNLL (175)
FLASFVEKTVL (32)	SPVVHQSL (80)	TLRSPGSSL (128)	GLRLIHVSL (176)
ILASHNLTV (33)	SPYRTPVL (81)	TVRGDVSSL (129)	GLRLLHVSL (177)
IQLTSVHLL (34)	SVFAGVVGV (82)	ALDPLLLRI (130)	LHNEKGLSL (178)
LELISFLPVL (35)	SYSPAHARL (83)	ISLIVTGLK (131)	LPSFSRPSGII (179)
NFCMLHQSI (36)	THGSEQLHL (84)	KILDVNLRI (132)	LSSRLPLGK (180)
PARPAGPL (37)	TQAPPNVVL (85)	ERVYIRASL (133)	MIGIKRLL (181)
PLPIVPAL (38)	VLVPYEPPQV (86)	ILDLESRY (134)	NLKKREIL (182)
SNLIRTGSH (39)	VSFPDVRKV (87)	KTFVQQKTL (135)	RMVAYLQQL (183)
VPAPAQAI (40)	VVFDKSDLAKY (88)	LYIKSLPAL (136)	SPARALPSL (184)
KGHGGPRSW (41)	YSHHSGLEY (89)	VLKEKNASL (137)	TVPGIQRY (185)
ITSSAVTTALK (42)	YYLDWIHHY (90)	LGISLTLKY (138)	VSRNYVLLI (186)
LLLPESPSI (43)	SVYKYLKAK (91)	DLLPKKLL (139)	LTVPLSVFW (187)
VILIPLPPK (44)	IYQFIMDRF (92)	HSLISIVYL (140)	KLNQAFLVL (188)
AVLLPKPPK (45)	GTLQGIRAW (93)	IAGALRSVL (141)	RLVSSTLLQK (189)
TQVSMAESI (46)	AQKVSVGQAA (94)	IGNPILRVL (142)	LPSHSLLI (190)
LNHLRTSI (47)	LYPSKLTHF (95)	IYAPHIRLS (143)	
NTSHLPLIY (48)	ATQNTIIGK (96)	LRSQILSY (144)	

[0016] .

[0017] 2. 항목 1에 있어서, 서열번호 97 내지 154에 제시된 아미노산 서열 중 하나를 포함하는, 백혈병 TAP.

[0018] 3. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*01:01 분자에 결합하고 그리고 아미노산 서열 NTSHLPLIY (서열번호 48), HTDDIENAKY(서열번호 67), YSHHSGLEY(서열번호 89), ILDLESRY(서열번호 134), VTDLLALTV(서열번호 151) 또는 LSDRQLSL(서열번호 164), 바람직하게는 ILDLESRY(서열번호 134) 또는 VTDLLALTV(서열번호 151)를 포함하는, 백혈병 TAP.

[0019] 4. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*02:01 분자에 결합하고 그리고 아미노산 서열 FLLEFKPVS (서열번호 7), LLSRGLLFRI(서열번호 11), LLDNILQSI(서열번호 27), FLASFVEKTVL(서열번호 32), ILASHNLTV(서열번호 33), IQLTSVHLL(서열번호 34), LELISFLPVL(서열번호 35), LLLPESPSI(서열번호 43), ALASHLIEA(서열번호 51), ALDDITIQL(서열번호 52), ALGNTVPAV(서열번호 53), ALLPAVPSL(서열번호 54), GLYYKLHNV(서열번호 61), HLLSETPQL(서열번호 65), KLLEKAFSI(서열번호 72), SLWGQPAEA(서열번호 77), SVFAGVVGV(서열번호 82), VLVPYEPPQV(서열번호 86), VLFGGKVSGA(서열번호 104), KLQDKEIGL(서열번호 108), TLNQGINVYI(서열번호 119), ALPVALPSL(서열번호 123), ALDPLLLRI(서열번호 130), KILDVNLRI(서열번호 132), SLLSGLLRA(서열번호 146), SLDLLPLSI(서열번호 150), ILLEEQSLI(서열번호 167), LTSISIRPV(서열번호 168), TISECPLLI(서열번호 169), ILLSNFSSL(서열번호 171), RMVAYLQQL(서열번호 183), 또는 KLNQAFLVL(서열번호 188), 바람직하게는 VLFGGKVSGA(서열번호 104), KLQDKEIGL(서열번호 108), TLNQGINVYI(서열번호 119), ALPVALPSL(서열번호 123), ALDPLLLRI(서열번호 130), KILDVNLRI(서열번호 132), SLLSGLLRA(서열번호 146) 또는 SLDLLPLSI(서열번호 150)를 포함하는, 백혈병 TAP.

[0020] 5. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*03:01 분자에 결합하고 그리고 아미노산 서열 RSASSATQVHK(서열번호 5), IVATGSLLK(서열번호 18), KIKNKTKNK(서열번호 19), KLLSLTIYK(서열번호 20), ITSSAVTTALK(서열번호 42), VILIPLPPK(서열번호 44), NVNRPLTMK(서열번호 74), SVYKYLKAK(서열번호 91), VVFPFPVNK(서열번호 105), ILFQNSALK(서열번호 113), TVIRIAIVNK(서열번호 126), ISLIVTGLK(서열번호 131), HVSDGSTALK(서열번호 159), IAYSVRALR(서열번호 160), LSSRLPLGK(서열번호 180) 또는 RLVSSTLLQK(서열번호 189), 바람직하게는 VVFPFPVNK(서열번호 105), ILFQNSALK(서열번호 113), TVIRIAIVNK(서열번호 126) 또는 ISLIVTGLK(서열번호 131)를 포함하는, 백혈병 TAP.

[0021] 6. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*11:01 분자에 결합하고 그리고 아미노산 서열 SASSATQVHK(서열번호 6), AVLLPKPPK(서열번호 45), ATQNTIIGK(서열번호 96), SLLIIPKKK(서열번호 106), SVQLLEQAIHK(서열번호 121), STFSLYLKK(서열번호 149) 또는 RTQITKVSLKK(서열번호 152), 바람직하게는 SLLIIPKKK(서열번호 106), SVQLLEQAIHK(서열번호 121), STFSLYLKK(서열번호 149) 또는 RTQITKVSLKK(서열번호 152)를 포함하는, 백혈병 TAP.

[0022] 7. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*24:02 분자에 결합하고 그리고 아미노산 서열 LYFLGHGSI (서열번호 13), NFCMLHQSI(서열번호 36), KFSNVTMLF(서열번호 71), IYQFIMDRF(서열번호 92), LYPSKLTHF(서열번호 95) 또는 RYLANKIHI(서열번호 145), 바람직하게는 RYLANKIHI(서열번호 145)를 포함하는, 백혈병 TAP.

[0023] 8. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*26:01 분자에 결합하고 그리고 아미노산 서열 ETTSQVRKY (서열번호 59) 또는 TVPGIQRY(서열번호 185)를 포함하는, 백혈병 TAP.

[0024] 9. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*29:02 분자에 결합하고 그리고 아미노산 서열 VVFDKSDLAKY(서열번호 88), FNVALNARY(서열번호 99) 또는 LGISLTLKY(서열번호 138), 바람직하게는 FNVALNARY (서열번호 99) 또는 LGISLTLKY(서열번호 138) 중 하나를 포함하는, 백혈병 TAP.

[0025] 10. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*30:01 분자에 결합하고 그리고 아미노산 서열 TSRLPKIQK(서열번호 26), LSWGYFLFK(서열번호 29) 또는 LSHPAPSSL(서열번호 165)을 포함하는, 백혈병 TAP.

[0026] 11. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-A*68:02 분자에 결합하고 그리고 아미노산 서열 NVSSHVHTV(서열번호 50) 또는 SSSPVRGPSV(서열번호 148), 바람직하게는 SSSPVRGPSV(서열번호 148)를 포함하는, 백혈병 TAP.

[0027] 12. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*07:02 분자에 결합하고 그리고 아미노산 서열 GPQVRGSI (서열번호 8), SPQSGPAL(서열번호 25), VPAPAQAI(서열번호 40), APAPPPVAV(서열번호 55), APDKKITL(서열번호 56), KPMPTKVVF(서열번호 73), SPADHRGYASL(서열번호 78), SPQSAAAEL(서열번호 79), SPVVHQSL(서열번호 80), SPYRTPVL(서열번호 81), PPRPLGAQV(서열번호 98), GPGSRESTL(서열번호 100), APGAAGQRL(서열번호 107), TPGRSTQAI(서열번호 110), APRGTAAL(서열번호 111), SPVVRVGL(서열번호 118), RPRGPRTAP(서열번호 120), TLRSPGSSL(서열번호 128), TVRGDVSSL(서열번호 129), LPSFSHFLLL(서열번호 157), PRGFLSAL(서열번호 161), IPLNPFSSL(서열번호 163), LPSFSRPSGII(서열번호 179) 또는 SPARALPSL(서열번호 184), 바람직하게는 PPRPLGAQV(서열번호 98), GPGSRESTL(서열번호 100), APGAAGQRL(서열번호 107), TPGRSTQAI(서열번호 110), APRGTAAL(서열번호 111), SPVVRVGL(서열번호 118), RPRGPRTAP(서열번호 120), TLRSPGSSL(서열번호 128) 또는 TVRGDVSSL(서열번호 129)을 포함하는, 백혈병 TAP.

[0028] 13. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*08:01 분자에 결합하고 그리고 아미노산 서열 SGKLRVAL (서열번호 4), NPLQLSLSI(서열번호 14), DLMLRESL(서열번호 15), IALYKQVL(서열번호 17), NILKKTVL(서열번호 21), NPKLKDIL(서열번호 22), NQKKVRIL(서열번호 23), RLEVRKVIL(서열번호 28), EGKIKRNI(서열번호 31), LNHLRTSI(서열번호 47), SIQRNLSL(서열번호 49), IPHQRSSL(서열번호 101), NLKEKKALF(서열번호 103), ILKKNISI(서열번호 114), VLKEKNASL(서열번호 137), DLLPKKLL(서열번호 139), SRIHLVVL(서열번호 147), QIKTKLLGSL(서열번호 156), TLKLKKIFF(서열번호 170), MIGIKRLL(서열번호 181) 또는 NLKKREIL(서열번호 182), 바람직하게는 IPHQRSSL(서열번호 101), NLKEKKALF(서열번호 103), ILKKNISI(서열번호 114), VLKEKNASL(서열번호 137), DLLPKKLL(서열번호 139) 또는 SRIHLVVL(서열번호 147)을 포함하는, 백혈병 TAP.

[0029] 14. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*14:01 분자에 결합하고 그리고 아미노산 서열 DRELRNLEL(서열번호 2), SNLIRTGSH(서열번호 39), DQVIRLAGL(서열번호 58), HQLYRASAL(서열번호 66), SLQILVSSL(서열번호 124), ERVYIRASL(서열번호 133), LYIKSLPAL(서열번호 136), IAGALRSVL(서열번호 141), ISSWLISSL(서열번호 162), DRGILRNLL(서열번호 175), GLRLIHVSL(서열번호 176) 또는 GLRLLHVSL(서열번호 177), 바람직하게는 SLQILVSSL(서열번호 124), ERVYIRASL(서열번호 133), LYIKSLPAL(서열번호 136) 또는 IAGALRSVL (서열번호 141)을 포함하는, 백혈병 TAP.

[0030] 15. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*15:01 분자에 결합하고 그리고 아미노산 서열 KIKVFSKVY(서열번호 10), AQMNLLQKY(서열번호 57), GQKPVILTY(서열번호 62) 또는 AQKVSVGQAA(서열번호 94)를 포함하는, 백혈병 TAP.

[0031] 16. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*27:05 분자에 결합하고 그리고 아미노산 서열 RQISVQASL(서열번호 1) 또는 LRSQILSY(서열번호 144), 바람직하게는 LRSQILSY(서열번호 144)를 포함하는, 백혈병 TAP.

[0032] 17. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*38:01 분자에 결합하고 그리고 아미노산 서열 TQVSMAESI(서열번호 46), HHLVETLKF(서열번호 64) 또는 THGSEQLHL(서열번호 84)을 포함하는, 백혈병 TAP.

[0033] 18. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*40:01 분자에 결합하고 그리고 아미노산 서열 REPYELTVPAL(서열번호 75) 또는 SEAEAAKNAL(서열번호 76)을 포함하는, 백혈병 TAP.

[0034] 19. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*44:03 분자에 결합하고 그리고 아미노산 서열 KEIFLELRL(서열번호 127)을 포함하는, 백혈병 TAP.

[0035] 20. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*51:01 분자에 결합하고 그리고 아미노산 서열 LPIASASLL(서열번호 12), PFPLVQVEPV(서열번호 24), PLPIVPAL(서열번호 38), IAAPILHV(서열번호 68), IPLAVRTI(서열번호 115), LPRNKPLL(서열번호 116) 또는 LPSHSLLI(서열번호 190), 바람직하게는 IPLAVRTI(서열번호 115) 또는 LPRNKPLL(서열번호 116)을 포함하는, 백혈병 TAP.

[0036] 21. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*57:01 분자에 결합하고 그리고 아미노산 서열 GARQQIHSW(서열번호 3), VTFKLSLF(서열번호 16), KGHGGPRSW(서열번호 41), GSLDFQRGW(서열번호 63), KAFPFHIIF(서열번호 69), GTLQGIRAW(서열번호 93), RTPKNYQHW(서열번호 122), ISNKVPKLF(서열번호 125), KTFVQQKTL(서열번호 135), ILRSPLKW(서열번호 153) 또는 LTVPLSVFW(서열번호 183), 바람직하게는 RTPKNYQHW(서열번호 122), ISNKVPKLF(서열번호 125), KTFVQQKTL(서열번호 135) 또는 ILRSPLKW(서열번호 153)를 포함하는, 백혈병 TAP.

[0037] 22. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-B*57:03 분자에 결합하고 그리고 아미노산 서열 GGSLIHPQW(서열번호 60) 또는 LGGAWKAVF(서열번호 172)를 포함하는, 백혈병 TAP.

[0038] 23. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*03:03 분자에 결합하고 그리고 아미노산 서열 PARPAGPL(서열번호 37), IASPIALL(서열번호 112) 또는 HSLISIVYL(서열번호 140), 바람직하게는 IASPIALL(서열번호 112) 또는 HSLISIVYL(서열번호 140)을 포함하는, 백혈병 TAP.

[0039] 24. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*05:01 분자에 결합하고 그리고 아미노산 서열 SLDLLPLSI(서열번호 150)를 포함하는, 백혈병 TAP.

[0040] 25. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*06:02 분자에 결합하고 그리고 아미노산 서열 IRMKAQAL(서열번호 9), KATEYVHSL(서열번호 70), VSFPDVRKV(서열번호 87), IGNPILRVL(서열번호 142), LSTGHLSTV(서열번호 154) 또는 LRKAVDPIL(서열번호 166), 바람직하게는 IGNPILRVL(서열번호 142) 또는 LSTGHLSTV(서열번호 154)를 포함하는, 백혈병 TAP.

[0041] 26. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*07:01 분자에 결합하고 그리고 아미노산 서열 IGNPILRVL(서열번호 142), IYAPHIRLS(서열번호 143), TVEEYLVNI(서열번호 155), LHNEKGLSL(서열번호 178) 또는 VSRNYVLLI(서열번호 186), 바람직하게는 IGNPILRVL(서열번호 142) 또는 IYAPHIRLS(서열번호 143)를 포함하는, 백혈병 TAP.

[0042] 27. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*07:02 분자에 결합하고 그리고 아미노산 서열 TILPRILTL(서열번호 30), SYSPAHARL(서열번호 83), TQAPPNVVL(서열번호 85), YYLDWIHHY(서열번호 90), SLREPQPAL(서열번호 109), PAPPHPAAL(서열번호 117) 또는 CLRIGPVTL(서열번호 158), 바람직하게는 SLREPQPAL(서열번호 109) 또는 PAPPHPAAL(서열번호 117)을 포함하는, 백혈병 TAP.

[0043] 28. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*08:02 분자에 결합하고 그리고 아미노산 서열 AQDIILQAV(서열번호 97), LTDRIYLTL(서열번호 102) 또는 AGDIIARLI(서열번호 174), 바람직하게는 AQDIILQAV(서열번호 97) 또는 LTDRIYLTL(서열번호 102)을 포함하는, 백혈병 TAP.

[0044] 29. 항목 1 또는 2에 있어서, 상기 백혈병 TAP는 HLA-C*12:03 분자에 결합하고 그리고 아미노산 서열 LSASHLSSL(서열번호 173)을 포함하는, 백혈병 TAP.

[0045] 30. 항목 1 내지 29 중 어느 하나에 있어서, 게놈의 비-단백질 코딩 영역에 존재하는 서열에 의해 인코딩되는, 백혈병 TAP.

[0046] 31. 항목 30에 있어서, 상기 게놈의 상기 비-단백질 코딩 영역은 비번역 전사된 영역(untranslated transcribed region, UTR)인, 백혈병 TAP.

[0047] 32. 항목 30에 있어서, 상기 게놈의 상기 비-단백질 코딩 영역은 인트론인, 백혈병 TAP.

[0048] 33. 항목 30에 있어서, 상기 게놈의 상기 비-단백질 코딩 영역은 유전자간 영역인, 백혈병 TAP.

[0049] 34. 항목 1 내지 33 중 어느 하나에 정의된 백혈병 TAP의 적어도 2가지를 포함하는 조합물.

[0050] 35. 항목 1 내지 33 중 어느 하나의 백혈병 TAP 또는 항목 34의 조합물을 인코딩하는, 핵산.

[0051] 36. 항목 35에 있어서, mRNA 또는 바이러스 벡터인, 핵산.

[0052] 37. 항목 1 내지 33 중 어느 하나의 백혈병 TAP, 항목 34의 조합물, 또는 항목 35 또는 36의 핵산을 포함하는, 리포솜.

[0053] 38. 항목 1 내지 33 중 어느 하나의 백혈병 TAP, 항목 34의 조합물, 항목 35 또는 36의 핵산, 또는 항목 37의 리포솜, 및 약제학적으로 허용 가능한 담체를 포함하는, 조성물.

[0054] 39. 항목 1 내지 33 중 어느 하나의 백혈병 TAP, 항목 34의 조합물, 항목 35 또는 36의 핵산, 항목 37의 리포솜, 또는 항목 38의 조성물 및 애주번트(adjuvant)를 포함하는, 백신.

[0055] 40. 항목 1 내지 33 중 어느 하나의 백혈병 TAP를 펩타이드 결합 홈에 포함하는, 단리된 주조직적합성 복합체(MHC) 클래스 I 분자.

[0056] 41. 항목 40에 있어서, 다량체의 형태인, 단리된 MHC 클래스 I 분자.

[0057] 42. 항목 41에 있어서, 상기 다량체는 사량체인, 단리된 MHC 클래스 I 분자.

[0058] 43. (i) 항목 1 내지 33 중 어느 하나의 백혈병 TAP, (ii) 항목 34의 조합물 또는 (iii) 항목 1 내지 33 중 어느 하나의 TAP 또는 항목 34의 조합물을 인코딩하는 뉴클레오타이드 서열을 포함하는 벡터를 포함하는, 단리된 세포.

[0059] 44. 항목 1 내지 33 중 어느 하나의 백혈병 TAP 또는 항목 34의 조합물을 펩타이드 결합 홈에 포함하는 주조직적합성 복합체(MHC) 클래스 I 분자를 표면에서 발현하는, 단리된 세포.

[0060] 45. 항목 44에 있어서, 항원-제시 세포(APC)인, 세포.

[0061] 46. 항목 45에 있어서, 상기 APC는 수지상 세포인, 세포.

[0062] 47. 항목 40 내지 42 중 어느 하나의 단리된 MHC 클래스 I 분자 및/또는 항목 44 내지 46 중 어느 하나의 세포의 표면에서 발현되는 MHC 클래스 I 분자를 특이적으로 인식하는, T-세포 수용체(TCR).

[0063] 48. 항목 47에 있어서, 상기 TCR은 서열번호 191 내지 219에 제시된 아미노산 서열에 제시된 아미노산 서열 중 하나를 포함하는 상보성 결정 영역 3(CDR3)을 포함하는 TCR베타(TCRβ) 사슬을 포함하는, TCR.

[0064] 49. 항목 47 또는 항목 48의 TCR을 세포 표면에 발현하는, 단리된 세포.

[0065] 50. 항목 49에 있어서, $CD8^+$ T 림프구인, 단리된 세포.

[0066] 51. 적어도 0.5%의 항목 49 또는 50에 정의된 바와 같은 단리된 세포를 포함하는, 세포 집단.

[0067] 52. 대상체에서 백혈병을 치료하는 방법으로서, 상기 대상체에게 유효량의, (i) 항목 1 내지 33 중 어느 하나의 백혈병 TAP; (ii) 항목 34의 조합물; (iii) 항목 35 또는 36의 핵산; (iv) 항목 37의 리포솜; (v) 항목 38의 조성물; (vi) 항목 39의 백신; (vii) 항목 43 내지 46, 49 및 50 중 어느 하나의 세포; 또는 (viii) 항목 51의 세포 집단을 투여하는 단계를 포함하는, 방법.

[0068] 53. 항목 52에 있어서, 상기 백혈병은 골수성 백혈병인, 방법.

[0069] 54. 항목 53에 있어서, 상기 골수성 백혈병은 급성 골수성 백혈병(AML)인, 방법.

[0070] 55. 항목 52 내지 54 중 어느 하나에 있어서, 상기 대상체에게 적어도 하나의 추가의 항종양 제제 또는 요법을 투여하는 단계를 더 포함하는, 방법.

[0071] 56. 항목 55에 있어서, 상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 방법.

[0072] 57. 대상체에서 백혈병을 치료하기 위한, (i) 항목 1 내지 33 중 어느 하나의 백혈병 TAP; (ii) 항목 34의 조합물; (iii) 항목 35 또는 36의 핵산; (iv) 항목 37의 리포솜; (v) 항목 38의 조성물; (vi) 항목 39의 백신; (vii) 항목 43 내지 46, 49 및 50 중 어느 하나의 세포; 또는 (viii) 항목 51의 세포 집단의 용도.

[0073] 58. 대상체에서 백혈병을 치료하기 위한 의약의 제조를 위한, (i) 항목 1 내지 33 중 어느 하나의 백혈병 TAP; (ii) 항목 34의 조합물; (iii) 항목 35 또는 36의 핵산; (iv) 항목 37의 리포솜; (v) 항목 38의 조성물; (vi) 항목 39의 백신; (vii) 항목 43 내지 46, 49 및 50 중 어느 하나의 세포; 또는 (viii) 항목 51의 세포 집단의 용도.

[0074] 59. 항목 57 또는 항목 58에 있어서, 상기 백혈병은 골수성 백혈병인, 용도.

[0075] 60. 항목 59에 있어서, 상기 골수성 백혈병은 급성 골수성 백혈병(AML)인, 용도.

[0076] 61. 항목 57 내지 60 중 어느 하나에 있어서, 적어도 하나의 추가의 항종양 제제 또는 요법의 사용을 더 포함하는, 용도.

[0077] 62. 항목 61에 있어서, 상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 방법.

[0078] 63. 대상체에서 백혈병의 치료에 사용하기 위한, (i) 항목 1 내지 33 중 어느 하나의 백혈병 TAP; (ii) 항목 34의 조합물; (iii) 항목 35 또는 36의 핵산; (iv) 항목 37의 리포솜; (v) 항목 38의 조성물; (vi) 항목 39의 백신; (vii) 항목 43 내지 46, 49 및 50 중 어느 하나의 세포; 또는 (viii) 항목 51의 세포 집단.

[0079] 64. 항목 63에 있어서, 상기 백혈병은 골수성 백혈병인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

[0080] 65. 항목 64에 있어서, 상기 골수성 백혈병은 급성 골수성 백혈병(AML)인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

[0081] 66. 항목 63 내지 65 중 어느 하나에 있어서, 적어도 하나의 추가의 항종양 제제 또는 요법과 조합하여 사용하기 위한, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

[0082] 67. 항목 66에 있어서, 상기 적어도 하나의 추가의 항종양 제제 또는 요법은 화학요법제, 면역요법, 면역관문저해제, 방사선요법 또는 수술인, 사용하기 위한 백혈병 TAP, 조합물, 핵산, 리포솜, 조성물, 백신, 세포 또는 세포 집단.

[0083] 본 발명의 다른 목적, 이점 및 특징은 첨부된 도면을 참조하여 단지 예로서 주어진 본 발명의 특정 실시형태의 하기의 비제한적인 설명을 읽으면 더욱 명백해질 것이다.

도면의 간단한 설명

[0084] 첨부 도면에서:

도 1A 내지 D는 조혈 선조체가 AML에서 TSA를 발견하기 위해 mTEC보다 더 양호한 대조군임을 나타내는 그래프이다. **도 1A**: 6개의 mTEC 샘플로부터 또는 6개의 MPC 샘플로부터 조합된 k-mer에 의해 19개 AML 시편의 각각의 k-mer 세트로부터의 k-mer 고갈(depletion)의 효능의 비교. 2개 미만의 K-MER 존재는 무시되었고 이 비교를 위하여 정규 모드에서 해파리 데이터베이스가 생성되었다. **도 1B**: **도 1A**에서 사용된 6개 mTEC 샘플 및 6개 MPC 샘플로부터의 k-mer와 모든 AML 시편의 조합된 k-mer 간의 중첩. **도 1A**에서 사용된 데이터베이스 구축을 위한 파라미터가 여기서 다시 적용되었다. **도 1C**: 표시된 조직으로부터의 정제된 세포 집단에서 발현된 단백질 코딩 유전자(TPM≥1)의 t-분산 확률 이웃 임베딩(t-distributed Stochastic Neighbor Embedding: t-SNE) 분석. **도 1D**: 패널 C를 플롯하는 데 사용되는 표시된 조직 및 세포 집단에서의 발현된 단백질 코딩 유전자(TPM≥1)의 총수의 비교. Pluri_stem: 다능성 줄기세포; Ery: 적혈구; Precu: 전구체; Lympho: 림프구; Granulo: 과립구; Mono: 단구성. 만 휘트니(Mann whitney) U 검정은 서로 다른 조직과 mTEC를 비교하는 데 사용되었고(****p<0.0001), 막대는 -기반 편차를 가진 평균을 나타낸다.

도 2는 MPC-기반 TSA 발견 접근법의 개략적 개요도를 도시한다. **(A)** mTEC k-mer 고갈에 기반한 TSA 발견을 위한 작업흐름의 개략적 개요도. **(B)** ERE-유래 MAP 발견을 위한 작업흐름의 개략적 개요도. **(C)** mTEC+MPC k-mer 고갈 TSA 발견 접근법을 위한 작업흐름의 개략적 개요도. **(D)** DKE 접근법을 위한 작업흐름의 개략적 개요도. 여기에서는 AML#1 샘플에 대한 작업흐름이 예시되어 있다. k-mer가 과발현된 것으로 간주하기 위하여 10의 배수 변화가 최소로 사용되었다(다른 필터도 적용되었으며, 방법을 참조한다). 3가지 다른 접근법에 대해서, RNA 시퀀싱을 수행하는 데 사용된 동일한 AML 샘플로부터 용출된 MAP의 MS 식별을 사용하기 전에 인 실리코 모든-프레임 번역된 콘티그(*in silico* all-frame translated contig)의 얻어진 데이터베이스를 개인화된 정규 프로테옴과 결부시켰다(concatenated).

도 3A 내지 H는 MPC-기반 접근법이 AML에서 대다수의 TSA^{hi}를 식별하는 것을 나타낸다. (검정으로 플로팅된) 분포의 σ)가 부여된다. **도 3A**: 각각의 AML 시편(n=19)에 대해서, TPM 발현에 기반한 10개의 상이한 군(십분위수)에서 분리된 전사체로부터 유래된 MAP의 비율. 십분위수 10은 최대 발현을 가진 전사체를 갖고, 십분위수 1

은 최저 발현을 가진 전사체를 갖는다. 박스는 분포의 중앙, 25번째 및 75번째 백분위수를 나타내고, 위스커(whisker)는 최소와 최대로 연장된다. **도 3B**: MS에 의해 식별된 AML 시편에서 이를 코딩할 수 있는 RNA-seq 리드의 총수의 대수 함수에서의 MAP의 누적 빈도의 정상 분포(도트). (검정으로 플로팅된) 분포의 평균(μ) 및 표준 편차(σ)가 부여된다. **도 3C**: 상이한 표시된 배수 변화(FC, 원래의 rphm×FC) 후 MAP를 생성하도록 RNA 서열에 대해서 (도 3B의 정상 분포 파라미터에 기반하여 계산된) 가능성. **도 3D**: 관심 MAP(MOI)를 TAA, HSA, TSA^{hi}로 분리하는 데 사용된 결정 트리. "정상 조직"은 모든 조직(GTEx, 정제된 조혈 세포 및 mTEC)을 지칭하고, 혈액/BM은 단지 정제된 조혈 세포만을 지칭한다. **도 3E**: 각각 표시된 프로테오게노믹 접근법에 의해 얻어진 MOI 계수의 비교. **도 3F**: 표시된 접근법들 간에 TSA^{hi} 동일성을 비교하는 벤 다이어그램. **도 3G**: 관찰된 체류 시간과 예측된 체류 시간(좌측) 또는 소수성 지수(우측) 간의 피어슨(Pearson) 상관관계. **도 3H**: 표시된 MAP의 Comet과의 성공적인 재식별의 중앙 및 사분위간 범위 빈도.

도 4A 내지 K는 TSA^{hi}가 인트론 번역으로부터 주로 유도되고 많은 환자 간에 공유되는 것을 나타낸다. **도 4A**: GTEx로부터의 전체 정상 조직(n = 이용 가능한 샘플에 따라서 12 내지 50), 정상의 분류된 조혈 세포 집단(n = 이용 가능한 샘플에 따라서 3 내지 16) 또는 mTEC(n=11)에서 각 식별된 TSA^{hi}의 평균 RNA 발현(rphm +1의 로그)을 나타내는 히트맵. 임상 시험에서 안전한 것으로 평가된 TAA가 또한 보고된다. Prec: 전구체. **도 4B**: 19개 AML 시편 및 MPC에서의 평균 rphm 발현(n=16)의 TSA^{hi} 배수 변화의 비교. 도트는 각각의 MOI를 나타내고, 박스는 분포의 중앙, 25번째 및 75번째 백분위수를 나타내고, 위스커는 최소와 최대로 연장된다. **도 4C**: 표시된 MOI가 생성되는 생물형(게놈 영역 또는 이벤트)의 분포. 엑손-인트론: 엑손-인트론 접합부와 중첩하는 펩타이드(보유된 인트론); ncRNA: 비-코딩 RNA; OoF 번역: 프레임밖(Out-of-frame) 번역. **도 4D**: 19개 AML 샘플에서 그리고 437명의 Leucegene 환자에서의 TSA^{hi} RNA 발현. **도 4E**: TSA^{hi}를 제시할 수 있는 HLA 알로타입에 의한 집단 범위(19개 AML 시편 대립유전자는 MHC클러스터에 의해 계산된 TSA^{hi} + 무차별 결합제(promiscuous binder)를 제시함). 이것은 IEDB 집단 범위 툴(www.iedb.org)로 계산되었다. 막대는 최대 6개의 알로타입을 보유하는 세계 집단 내에서의 개체의 빈도를 표시하고(x-축), 집단 범위의 누적 백분율은 도트로 표시된다. **도 4F**: TSA^{hi} RNA 발현에 기반한 Leucegene 코호트(rphm≥2인 경우 발현된 것으로 간주됨), 환자의 HLA 대립유전자(OptiType) 및 무차별 결합제에서의 $HLA\text{-}TSA^{hi}$ 복합체 분포. 높은(상위 사분위수)과 낮은(모든 다른 환자) TSA^{hi} 발현자(expressor) 간의 차이가 도시되어 있다. **도 4G**: 진단 및 재발에서 Leucegene 환자에서의 $_{pred}HLA\text{-}TSA^{hi}$ 복합체 개수. **도 4H**: 진단 시 및 재발에서 15명의 환자로부터 쌍별-정제된 AML 모세포의 HLA 대립유전자에 의해 제시될 수 있는 TSA^{hi}의 RNA 발현((Toffalori *et al.*, 2019)로부터의 데이터). 윌콕슨 일치-쌍 부호 순위 검정(Wilcoxon matched-pairs signed rank test)으로 행해진 비교. **도 4I**: 어디에선가(Corces *et al.*, 2016) 보고된, 분류된 모세포(n=12) 또는 백혈병 줄기세포(LSC, n=8)에서 샘플간 공유하는 TSA^{hi}의 비교(rphm>0인 경우 발현된 것으로 간주됨). **도 4J**: 도 4I에 나타낸 샘플에서의 HLA-ABC 분자의 RNA 발현. 평균 +SD가 표시되어 있다. **도 4K**: 표시된 LSC 시그니처 유전자 세트(Eppert *et al*. 2011)에 대해서 TSA^{hi}의 중앙 개수 이상을 발현하는 Leucegene 환자(rphm>0)(n=207) 대 기타(n=230)를 비교하는 GSEA 분석. NES, 정규화된 농축 스코어(normalized enrichment score).

도 5A 내지 F는 다수의 TSA^{hi}의 제시가 더 양호한 생존율과 상관이 있는 것을 나타낸다. **도 5A**: 높은(고) 개수(n=98, **도 5B**에서의 상위 사분위수) 대 낮은(저) 개수(n=275, 모든 다른 환자)의 $HLA\text{-}TSA^{hi}$ 복합체를 발현하는 Leucegene 환자 간의 카플란-마이어 생존 분석(Kaplan-Meier survival analysis). 통계학적 유의성은 로그 순위법(로그-순위 test)에 의해 결정되었다. **도 5B**: 5-년 전체 생존율의 다변수 분석에 대한 Forest 플롯. HR, 조정된 위험 비율; CI, 신뢰구간; adv, 유해; fav, 양호; int, 중간. NPM1/FLT3 상호작용 = $NPM1^{mut}$ 및 *FLT3*-ITD 둘 다의 존재. **도 5C**: (A)에서 수행된 분석으로부터의 표시된 수의 TSA^{hi}의 제거 후 산출된 로그-순위(로그-순위) p-값; 1000 순열이 각 수에 대해서 이루어졌고, 평균 +SD가 보고되어 있다. **도 5D**: 도 5C에서 얻어진 유의한 p-값의 백분율. **도 5E**: 도 5A에서의 분석으로부터의 각 TSA^{hi}의 대안적인 제거 후에 재산출된 로그-순위 p-값의 비교. **도 5F**: 도 5A에서의 분석으로부터의 각 HLA 대립유전자의 대안적인 제거 후에 재산출된 로그-순위 p-값의 비교.

도6A 내지 O는 TSA^{hi} 제시가 세포독성 T 세포 반응을 촉발시키는 것을 나타낸다. **도 6A**: 흉선기질 세포로부터의 MAP 및 HIV MAP의 MOI 간의 면역원성 스코어(Repitope)의 비교. **도 6B**: 5112개의 비-면역원성 MAP 및 1411개의 면역원성 MAP(IEDB로부터 입수되고 문헌(Ogishi and Yotsuyanagi, 2019)에서 엄선됨)의 MOI에 대한 11개의 입수 가능한 mTEC 샘플에 대한 평균 RNA 발현의 중앙 및 사분위간 범위. **도 6C**: 표시된 펩타이드로 펄싱된 DC에 의한 자극 후의 건강한 PBMC의 IFN-γ ELISpot 검정. 2개의 독립적인 실험으로부터의 결과가 조합되었다. **도 6D**: 표시된 TSA^{hi}(단일 공여체)의 ELISpot 검정. **도 6E**: 표시된 펩타이드의 존재에서 확장된 T 세포의 사이토카인 분비의 유세포분석. **도 6F**: 표시된 펩타이드의 존재 하에 확장된 T 세포 간의 표시된 덱스트라머(dextramer) 빈도의 대표적인 유세포분석 플롯. **도 6G**: FEST 검정: TSA^{hi}의 3개의 상이한 풀(pool)(5개 펩타이드/풀)에 의한 자극 10일 후의 유의한 T-세포 클론형의 확장. **도 6H**: 표시된 $_{pred}$HLA-MOI의 높은 대 낮은 계수치를 갖는 Leucegene 환자(도 4F 및 도 11D과 관련됨)에서의 천개의 TCR 리드당 TCR CDR3(CPK, 클론형 다양성의 척도로서). **도 6I**: Leucegene에서 TSA^{hi} (n = 66-164/군)에 대해서 반응하도록 ERGO에 의해 예측된 클론형의 빈도. **도 6J**: Leucegene에서 TAA(n = 74-207/군)에 대해서 반응하도록 ERGO에 의해 예측된 클론형의 빈도. **도 6K**: $_{pred}$제시된 MOI의 수로 정규화된(I 및 J와 관련됨) 모든 항-MOIs 클론형 중 고려된 샘플에서 $_{pred}$제시된 MOI를 인식하는 클론형의 빈도. 항-$_{pres}$MOI 클론형 계수치 = 0을 가진 환자는 무시되었다. **도 6L**: *CD8A* 및 *CD8B* 유전자의 RNA 발현과 Leucegene에서 2 rphm 초과를 발현하는 TSA^{hi}의 수 간의 상관관계. **도 6M**: *CD8A* 및 *CD8B* 유전자의 RNA 발현과 Leucegene에서 $_{pred}$HLA-TSA^{hi}의 개수 간의 상관관계 **도 6N**: 정규화된 TSA^{hi} $_{pred}$제시가 중앙보다 위 대 중앙보다 아래였던 환자를 비교한 유전자 발현량 차이 분석의 화산(Volcano) 플롯. 도트는 중앙보다 위족의 환자에서 상향 조절된 유전자를 나타낸다. **도 6O**: 도 6N에서의 상향 조절된 유전자의 GO 용어 분석.

도 7A 내지 G는 TSA^{hi} 발현이 면역편집된, AML 유발 돌연변이 및 후성적 비정상과 연관된 것을 나타낸다. **도 7A**: 전체 Leucegene 코호트(n=437)에 걸쳐서 표시된 유전자의 발현과 HE-TSA^{hi}의 수 간의 피어슨 상관관계. *HLA-A*, *-B* 및 *-C*의 발현값은 제1 패널에 대해 합산되었다. **도 7B**: HE-TSA^{hi}의 중앙값 이상을 발현하는 Leucegene 환자 대 *NPM1* 돌연변이 상태의 함수로서 계층화된 기타의 것 간의 *PD-L1*(*CD274*) 유전자 발현의 비교. **도 7C**: HE-TSA^{hi} 수와 역상관 관계에 있는 유전자 중 GO 용어 농축의 네트워크 분석. 노드 크기는 유전자 세트 크기에 비례한다. **도 7D**: HE-TSA^{hi} 수와 양성적 상관관계에 있는 유전자 중 GO 용어 농축의 네트워크 분석. **도 7E**: HE-TSA^{hi}의 중앙값 수 이상을 발현하는 환자 대 표시된 유전자에 대한 WT 및 돌연변이 환자 중 기타의 것의 비교. Fisher 정확 검정으로 확립된 통계학적 유의성(**p<0.01, ****p<0.0001). **도7F**: *NPM1*, *FLT3* 또는 *DNMT3A* 중 하나에서 0 내지 3개의 돌연변이를 가진 환자들 간의 HE-TSA^{hi} 수의 비교. **도 7G**: Leucegene 환자(n=437, 열(column))에 대해서 IRFinder에 의해 결정된 바와 같은 인트론 보유 비율에 대한 무감독 공통 클러스터(unsupervised consensus clustering). 행(row)은 계층적으로 클러스터링된 공통 클러스터링에 대한 최고 가변성 및 유의성의 1211개의 최상의 등급 인트론을 나타낸다. 환자의 FAB 유형은 표시된 공통 클러스터와 유의한 연관에 대해서 표시된 p-값(Fisher 정확 검정, *p<0.05, **p<0.01, ***p<0,001, ****p<0.0001)과 함께 히트맵 아래에 표시된다.

도 8A는 k-mer 발생의 개념의 예시이다.

도 8B는 샘플 05H143에서의 발생 함수에서 k-mer 빈도 분포의 예를 묘사하는 그래프이다.

도 8C는 mTEC 단독과 mTEC+MPC k-mer 고갈 접근법 간에 사용된 역치 발생의 비교를 묘사하는 그래프이다(각 도트는 상이한 AML 샘플이다).

도 8D는 mTEC 또는 mTEC+MPC k-mer의 고갈 후에 얻어진 모든 19개의 AML 시편으로부터 얻어진 고유한 k-mer의 조합 간의 k-mer 동일성의 중첩을 묘사하는 그래프이다.

도 9A는 k-mer 발현량 차이 분석 및 MS 데이터베이스 구축의 상세를 제공하는 개략도이다. 샘플 AML#1에 대한 MS 데이터베이스의 구축이 예로서 제시된다. 배수 변화인 FC는 k-mer 발현량 차이 분석 및 MS 데이터베이스 구축의 상세를 나타내는 다이어그램이다. 샘플 AML#1에 대한 MS 데이터베이스의 구축이 예로서 제시된다. FC, 배수 변화.

도 9B는 정규 펩타이드 식별의 누적 수(단독(Canon.)으로 또는 4개의 표시된 접근법에서의 콘티그 서열과 결부된, 개인화된 정규 프로테옴으로부터 유도된 펩타이드) 대 평균 데이터베이스 크기(라인)를 묘사하는 그래프이다.

도 9C는 각 접근법에 기반하여 식별된 정규 펩타이드의 동일성의 중첩을 정규 개인화된 프로테옴 단독에 기반하여 식별된 펩타이드와 비교하는 벤 다이어그램을 묘사한다.

도 10A는 각 TSA-식별 접근법에 의해 식별된 관심 MHC-I-연관 펩타이드(MAP)(MOI)의 비율의 비교를 묘사하는 그래프이다.

도 10B는 mTEC + MPC k-mer 고갈로 또는 k-mer 발현량 차이 접근법으로 식별된 TSA^{hi}를 발현하는 (rphm > 0) (본 연구에서 TSA를 식별하는 데 사용된 19개 중) AML 시편의 총수의 비교를 묘사하는 그래프이다.

도 11A는 Leucegene 코호트(n=437)에서 2 rphm과 동등하거나 더 높은 RNA 발현을 갖는 TSA^{hi} 수의 분보를 나타내는 그래프이다.

도 11B는 2 rphm와 동등 또는 그보다 높은 수준(좌측 패널에서의 분포의 상위 사분위수)에서 발현된 높은 수의 TSA^{hi}를 제시하는 Leucegene 코호트의 환자(진단에서 서열화되고 생존 데이터가 입수 가능하였던 n=372) 대 낮은 수준을 제시하는 자들(코호트의 나머지) 간의 생존율 비교를 나타내는 그래프이다.

도 11C 내지 E는 HSA(**도 11C**), TAA(**도 11D**) 및 TSA^{lo}(**도 11E**)에 대해서 각 환자의 HLA 대립유전자의 RNA 발현(rphm≥2인 경우 발현된 것으로 간주됨) 및 무차별 결합제 예측(옵티타입(optitype) 및 MHC클러스터)에 기반하여 얻어진 전체 Leucegene 코호트에 걸친 HLA-MOI 복합체 분포를 나타내는 그래프이다.

도 11F 내지 H는 HSA(**도 11F**), TAA(**도 11G**) 및 TSA^{lo}(**도 11H**)에 대해서 높은 수의 HLA-MOI 복합체(최상의 패널에서 분포의 상위 사분위수)를 제시하는 Leucegene 코호트의 환자(진단에서 서열화되고 생존 데이터가 입수 가능하였던 n=372) 대 낮은 수준을 제시하는 자들(코호트의 나머지) 간의 생존율 비교를 나타내는 그래프이다.

도 12A는 전체 Leucegene 코호트(n=437)에 걸쳐서 $HE\text{-}TSA^{hi}$의 수와 표시된 유전자의 발현 간의 피어슨 상관관계를 묘사한다.

도 12B는 높은 $_{pred}$제시 수준의 TSA^{hi}를 가진 환자 대 나머지 환자(도 4F와 관련)에서 표시된 유전자 발현의 비교를 나타내는 그래프를 묘사한다.

도 12C는 (IRFinder로 분석되고 10% 초과의 전사체에서 보유된 경우 보유된 것으로 정의된) Leucegene 코호트에서의 보유된 인트론과 ZNF445의 발현 간의 피어슨 상관관계를 묘사한다.

도 12D는 $HE\text{-}TSA^{hi}$의 중앙값 수 이상을 발현하는 환자 대 표시된 유전자에 대한 WT 및 돌연변이 환자 중 기타의 것의 비교를 나타내는 그래프이다. 통계학적 유의성은 Fisher 정확 검정으로 확립되었다.

도 12E는 $HE\text{-}TSA^{hi}$의 중앙값 수 이상을 발현하는 환자 대 allo-HSCT를 받거나 받지 않은 환자들 중 기타의 것의 비교를 나타내는 그래프이다. 통계학적 유의성은 Fisher 정확 검정으로 확립되었다.

도 12F는 전체 Leucegene 코호트에 걸쳐서 중앙 $HE\text{-}TSA^{hi}$ 계수치보다 높거나 또는 낮은 $HE\text{-}TSA^{hi}$ 계수치를 갖는 환자의 FAB 유형 분포를 나타내는 그래프이다.

도 12G는 전체 Leucegene 코호트에 걸쳐서 중앙 $HE\text{-}TSA^{hi}$ 계수치보다 높거나 또는 낮은 $HE\text{-}TSA^{hi}$ 계수치를 갖는 환자의 WHO 2008 분류 분포를 나타내는 그래프이다.

도 12H는 전체 Leucegene 코호트에 걸쳐서 중앙 $HE\text{-}TSA^{hi}$ 계수치보다 높거나 또는 낮은 $HE\text{-}TSA^{hi}$ 계수치를 갖는 환자의 세포유전학 프로파일 분포를 나타내는 그래프이다.

발명을 실시하기 위한 구체적인 내용

[0085] 본 명세서에서 사용되는 유전학, 분자 생물학, 생화학 및 핵산의 용어 및 기호는, 예컨대, 문헌[Kornberg and

Baker, DNA Replication, Second Edition (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Second Edition (Worth Publishers, New York, 1975); Strachan and Read, Human Molecular Genetics, Second Edition (Wiley-Liss, New York, 1999); Eckstein, editor, Oligonucleotides and Analogs: A Practical Approach (Oxford University Press, New York, 1991); Gait, editor, Oligonucleotide Synthesis: A Practical Approach (IRL Press, Oxford, 1984)] 등과 같은 당해 분야의 표준 논문 및 텍스트의 것들을 따른다. 모든 용어는 관련 분야에서 확립된 이들의 전형적인 의미를 갖는 것으로 이해되어야 한다.

[0086] 본 명세서에서 단수 표현의 관사는 관사의 문법상 목적어 중 하나 또는 하나 초과(즉, 적어도 하나)를 지칭하는 것으로 사용된다. 예로서, "요소"는 하나의 요소 또는 하나 초과의 요소를 의미한다. 본 명세서 전반에 걸쳐서, 문맥상 달리 요구되지 않는 한, "포함하다(comprise)", "포함한다" 및 "포함하는"이라는 단어는 언급된 단계 또는 요소, 또는 단계들 또는 요소들의 군을 포함하지만, 임의의 다른 단계 또는 요소, 또는 단계들 또는 요소들의 군을 배제하는 것이 아님을 암시하는 것으로 이해될 것이다.

[0087] 본 명세서에서 달리 나타내지 않는 한, 본 명세서에서 값을 범위의 언급은 단지 이러한 범위 내에 포함되는 각각의 개별 값을 개발적으로 지칭하는 속기 방법으로서 역할하는 것으로 의도되며, 각각의 개별 값은 마치 본 명세서에 개별적으로 언급된 것과 같이 본 명세서에 포함된다. 이러한 범위 내의 값의 모든 부분집합은 또한 마치 본 명세서에 개별적으로 언급된 것처럼 본 명세서에 포함된다.

[0088] 본 명세서에 기술된 모든 방법은, 본 명세서에서 달리 나타내지 않는 한, 또는 다르게는 문맥상 명백하게 모순되지 않는 한, 임의의 적합한 순서로 수행될 수 있다.

[0089] 본 명세서에 제공된 임의의 그리고 모든 예, 또는 예시적인 표현(예를 들어, "예컨대")의 사용은, 단지 본 발명을 더욱 잘 설명하기 위해 의도된 것이며, 달리 청구되지 않는 한, 본 발명의 범주를 제한하지 않는다.

[0090] 본 명세서에서 어떤 표현도 임의의 비청구된 요소를 본 발명의 실시에 필수적인 것으로서 명시하는 것으로 해석되어서는 안 된다.

[0091] 본 명세서에서, "약"이라는 용어는 그의 통상적인 의미를 갖는다. "약"이라는 용어는 값이 그 값을 측정하기 위해 사용되는 디바이스 또는 방법에 대한 고유한 오차 변동을 포함하거나, 또는 언급된 값에 가까운 값, 예를 들어, 언급된 값(또는 값의 범위)의 10% 또는 5% 이내를 포함한다는 것을 나타내기 위해 사용된다.

[0092] 본 명세서에 기재된 연구에서, 본 발명자들은 프로테오게노믹-기반 접근법을 사용하여 19개의 AML 시편으로부터 TSA 후보물질을 식별하였다. 이러한 TSA의 많은 부분은 비-엑손 서열(예컨대, 인트론 및 유전자간 서열)과 같은 정상 조직에서는 발현되지 않는 비정상적으로 발현된 돌연변이되지 않은 게놈 서열로부터 유래되었다. 이러한 AML TSA 후보물질의 발현은 후성유전적 변형자(예컨대, *DNMT3A*)의 돌연변이와 그리고 게놈 각인의 조절자인 *ZNF445*의 발현과 상관관계가 있는 것으로 나타났다. 또한 AML TSA 후보물질은 환자들 간에 고도로 공유되고, 모세포와 백혈병 줄기세포 둘 다에서 발현되며, 이들의 HLA 제시는 면역 편집 및 더 나은 전체 생존의 마커와 관련이 있는 것으로 제시되어 있다. 따라서, 본 명세서에서 식별된 신규한 AML TSA 후보물질은 백혈병 T-세포 기반 면역요법에 유용할 수 있다.

[0093] 따라서, 하나의 양상에서, 본 개시내용은 하기 아미노산 서열 중 하나를 포함하거나 이로 이루어진 백혈병 TAP (또는 백혈병 종양-특이적 펩타이드)에 관한 것이다:

[0094]

서열 (서열번호)	서열 (서열번호)	서열 (서열번호)	서열 (서열번호)
RQISVQASL (1)	SIQRNLSL (49)	AQDIILQAV (97)	RYLANKIHI (145)
DRELRNLEL (2)	NVSSHVHTV (50)	PPRPLGAQV (98)	SLLSGLLRA (146)
GARQQIHSW (3)	ALASHLIEA (51)	FNVALNARY (99)	SRIHLVVL (147)
SGKLRVAL (4)	ALDDITIQL (52)	GPGSRESTL (100)	SSSPVRGPSV (148)
RSASSATQVHK (5)	ALGNTVPAV (53)	IPHQRSSL (101)	STFSLYLKK (149)
SASSATQVHK (6)	ALLPAVPSL (54)	LTDRIYLTL (102)	SLDLLPLSI (150)
FLLEFKPVS (7)	APAPPPVAV (55)	NLKEKKALF (103)	VTDLLALTV (151)
GPQVRGSI (8)	APDKKITL (56)	VLFGGKVSGA (104)	RTQITKVSLKK (152)
IRMKAQAL (9)	AQMNLLQKY (57)	VVFPFPVNK (105)	ILRSPLKW (153)
KIKVFSKVY (10)	DQVIRLAGL (58)	SLLIIPKKK (106)	LSTGHLSTV (154)
LLSRGLLFRI (11)	ETTSQVRKY (59)	APGAAGQRL (107)	TVEEYLVNI (155)
LPIASASLL (12)	GGSLIHPQW (60)	KLQDKEIGL (108)	QIKTKLLGSL (156)
LYFLGHGSI (13)	GLYYKLHNV (61)	SLREPQPAL (109)	LPSFSHFLLL (157)
NPLQLSLSI (14)	GQKPVILTY (62)	TPGRSTQAI (110)	CLRIGPVTL (158)
DLMLRESL (15)	GSLDFQRGW (63)	APRGTAAL (111)	HVSDGSTALK (159)
VTFKLSLF (16)	HHLVETLKF (64)	IASPIALL (112)	IAYSVRALR (160)
IALYKQVL (17)	HLLSETPQL (65)	ILFQNSALK (113)	PRGFLSAL (161)
IVATGSLLK (18)	HQLYRASAL (66)	ILKKNISI (114)	ISSWLISSL (162)
KIKNKTKNK (19)	HTDDIENAKY (67)	IPLAVRTI (115)	IPLNPFSSL (163)
KLLSLTIYK (20)	IAAPILHV (68)	LPRNKPLL (116)	LSDRQLSL (164)
NILKKTVL (21)	KAFPFHIIF (69)	PAPPHPAAL (117)	LSHPAPSSL (165)
NPKLKDIL (22)	KATEYVHSL (70)	SPVVRVGL (118)	LRKAVDPIL (166)

[0095]

NQKKVRIL (23)	KFSNVTMLF (71)	TLNQGINVYI (119)	ILLEEQSLI (167)
PFPLVQVEPV (24)	KLLEKAFSI (72)	RPRGPRTAP (120)	LTSISIRPV (168)
SPQSGPAL (25)	KPMPTKVVF (73)	SVQLLEQAIHK (121)	TISECPLLI (169)
TSRLPKIQK (26)	NVNRPLTMK (74)	RTPKNYQHW (122)	TLKLKKIFF (170)
LLDNILQSI (27)	REPYELTVPAL (75)	ALPVALPSL (123)	ILLSNFSSL (171)
RLEVRKVIL (28)	SEAEAAKNAL (76)	SLQILVSSL (124)	LGGAWKAVF (172)
LSWGYFLFK (29)	SLWGQPAEA (77)	ISNKVPKLF (125)	LSASHLSSL (173)
TILPRILTL (30)	SPADHRGYASL (78)	TVIRIAIVNK (126)	AGDIIARLI (174)
EGKIKRNI (31)	SPQSAAAEL (79)	KEIFLELRL (127)	DRGILRNLL (175)
FLASFVEKTVL (32)	SPVVHQSL (80)	TLRSPGSSL (128)	GLRLIHVSL (176)
ILASHNLTV (33)	SPYRTPVL (81)	TVRGDVSSL (129)	GLRLLHVSL (177)
IQLTSVHLL (34)	SVFAGVVGV (82)	ALDPLLLRI (130)	LHNEKGLSL (178)
LELISFLPVL (35)	SYSPAHARL (83)	ISLIVTGLK (131)	LPSFSRPSGII (179)
NFCMLHQSI (36)	THGSEQLHL (84)	KILDVNLRI (132)	LSSRLPLGK (180)
PARPAGPL (37)	TQAPPNVVL (85)	ERVYIRASL (133)	MIGIKRLL (181)
PLPIVPAL (38)	VLVPYEPPQV (86)	ILDLESRY (134)	NLKKREIL (182)
SNLIRTGSH (39)	VSFPDVRKV (87)	KTFVQQKTL (135)	RMVAYLQQL (183)
VPAPAQAI (40)	VVFDKSDLAKY (88)	LYIKSLPAL (136)	SPARALPSL (184)
KGHGGPRSW (41)	YSHHSGLEY (89)	VLKEKNASL (137)	TVPGIQRY (185)
ITSSAVTTALK (42)	YYLDWIHHY (90)	LGISLTLKY (138)	VSRNYVLLI (186)
LLLPESPSI (43)	SVYKYLKAK (91)	DLLPKKLL (139)	LTVPLSVFW (187)
VILIPLPPK (44)	IYQFIMDRF (92)	HSLISIVYL (140)	KLNQAFLVL (188)
AVLLPKPPK (45)	GTLQGIRAW (93)	IAGALRSVL (141)	RLVSSTLLQK (189)
TQVSMAESI (46)	AQKVSVGQAA (94)	IGNPILRVL (142)	LPSHSLLI (190)
LNHLRTSI (47)	LYPSKLTHF (95)	IYAPHIRLS (143)	
NTSHLPLIY (48)	ATQNTIIGK (96)	LRSQILSY (144)	

[0096] 일반적으로, HLA 클래스 I의 맥락에서 제시되는 TAP와 같은 펩타이드는 약 7 또는 8 내지 약 15개, 또는 바람직하게는, 8 내지 14개의 아미노산 잔기의 길이로 다양하다. 본 개시내용의 방법의 일부 실시형태에서, 본 명세서에 정의된 TAP 서열을 포함하는 더 긴 펩타이드는 항원 제시 세포(APC)와 같은 세포에 인공적으로 로딩되고, 세포에 의해 프로세싱되고, TAP는 APC의 표면에서 MHC 클래스 I 분자에 의해 제시된다. 이 방법에서, 15개 아미노산 잔기보다 더 긴 펩타이드/폴리펩타이드가 APC에 로딩될 수 있고, APC 사이토졸에서 프로테아제에 의해 프로세싱되고, 이로써 제시를 위한 본 명세서에 정의된 바와 같은 상응하는 TAP를 제공한다. 일부 실시형태에서, 본 명세서에 정의된 TAP를 생성하는 데 사용되는 전구체 펩타이드/폴리펩타이드는, 예를 들어, 1000, 500, 400, 300, 200, 150, 100, 75, 50, 45, 40, 35, 30, 25, 20 또는 15개 이하의 아미노산이다. 따라서, 본 명세서에 기재된 TAP를 사용하는 모든 방법 및 프로세스는, 세포(APC)에 의한 프로세싱 후에 "최종" 8-14 TAP의 제시를 유도하기 위해, 더 긴 펩타이드 또는 폴리펩타이드(천연 단백질 포함), 즉, 종양 항원 전구체 펩타이드/폴리펩타이드를 사용하는 것을 포함한다. 일부 실시형태에서, 본 명세서에서 언급된 TAP는, HLA 클래스 I 분자에서의 직접 피트를 위해 충분히 작은, 약 8 내지 14, 8 내지 13, 또는 8 내지 12개 아미노산 길이(예컨대, 8, 9, 10, 11, 12 또는 13개 아미노산 길이)이다. 실시형태에서, TAP는 20개 이하의 아미노산, 바람직하게는 15개 이하의 아미노산, 더 바람직하게는 14개 이하의 아미노산을 포함한다. 실시형태에서, TAP는 적어도 7개의 아미노산, 바람직하게는 적어도 8개 이하의 아미노산, 더 바람직하게는 적어도 9개의 아미노산을 포함한다.

[0097] 본 명세서에서 사용되는 바와 같은 용어 "아미노산"은 펩타이드 화학에서 TAP의 합성 유사체를 제조하는 데 사용되는 천연형 아미노산뿐만 아니라, 다른 아미노산(예컨대, 천연형 아미노산, 비천연형 아미노산, 핵산 서열에 의해 인코딩되지 않은 아미노산 등)의 L- 및 D-이성질체를 둘 다 포함한다. 천연형 아미노산의 예는 글리신, 알

라닌, 발린, 류신, 아이소류신, 세린, 트레오닌 등이다. 기타 아미노산은, 예를 들어, 비유전적으로 인코딩된 형태의 아미노산뿐만 아니라, L-아미노산의 보존적 치환을 포함한다. 천연형 비유전적으로 인코딩된 아미노산은, 예를 들어, 베타-알라닌, 3-아미노-프로피온산, 2,3-다이아미노프로피온산, 알파-아미노아이소부티르산(Aib), 4-아미노-부티르산, *N*-메틸글리신(사르코신), 하이드록시프롤린, 오르니틴(예컨대, L-오르니틴), 시트룰린, *t*-부틸알라닌, *t*-부틸글리신, *N*-메틸아이소류신, 페닐글리신, 사이클로헥실알라닌, 노르류신(Nle), 노르발린, 2-나프틸알라닌, 피리딜알라닌, 3-벤조티에닐 알라닌, 4-클로로페닐알라닌, 2-플루오로페닐알라닌, 3-플루오로페닐알라닌, 4-플루오로페닐알라닌, 페니실라민, 1,2,3,4-테트라하이드로-아이소퀴놀린-3-카복실산, 베타-2-티에닐알라닌, 메티오닌 설폭사이드, L-호모아르기닌(Hoarg), N-아세틸 라이신, 2-아미노 부티르산, 2-아미노 부티르산, 2,4,-다이아미노부티르산(D- 또는 L-), *p*-아미노페닐알라닌, *N*-메틸발린, 호모시스테인, 호모세린(HoSer), 시스테인산, 엡실론-아미노 헥산산, 델타-아미노 발레르산 또는 2,3-다이아미노부티르산(D- 또는 L-) 등을 포함한다. 이들 아미노산은 생화학/펩타이드 화학 분야에 잘 알려져 있다. 실시형태에서, TAP는 천연형 아미노산만을 포함한다.

[0098] 실시형태에서, 본 명세서에 기재된 TAP는, 본 명세서에서 언급된 서열과 비교하여, 기능적으로 등가인 아미노산 잔기의 치환을 함유하는 변경된 서열을 갖는 펩타이드를 포함한다. 예를 들어, 서열 내의 하나 이상의 아미노산 잔기는 기능적 등가물로서 작용하는, 유사한 극성의 (유사한 물리-화학적 특성을 갖는) 또 다른 아미노산에 의해 치환됨으로써 침묵 변경을 일으킬 수 있다. 서열 내의 아미노산에 대한 치환은 그 아미노산이 속하는 부류의 다른 구성원으로부터 선택될 수 있다. 예를 들어, 양으로 하전된(염기성) 아미노산은 아르기닌, 라이신 및 히스티딘(뿐만 아니라, 호모아르기닌 및 오르니틴)을 포함한다. 비극성(소수성) 아미노산은 류신, 아이소류신, 알라닌, 페닐알라닌, 발린, 프롤린, 트립토판 및 메티오닌을 포함한다. 비하전된 극성 아미노산은 세린, 트레오닌, 시스테인, 티로신, 아스파라긴 및 글루타민을 포함한다. 음으로 하전된(산성) 아미노산은 글루탐산 및 아스파르트산을 포함한다. 아미노산 글리신은 비극성 아미노산 패밀리 또는 비하전된(중성) 극성 아미노산 패밀리에 포함될 수 있다. 아미노산 패밀리 내에서 이루어진 치환은 일반적으로 보존적 치환인 것으로 이해된다. 본 명세서에서 언급된 TAP는 모든 L-아미노산, 모든 D-아미노산 또는 L- 아미노산과 D-아미노산의 혼합물을 포함할 수 있다. 실시형태에서, 본 명세서에서 언급된 TAP는 모든 L-아미노산을 포함한다.

[0099] 실시형태에서, 서열번호 1 내지 190, 바람직하게는 서열번호 97 내지 154의 서열 중 하나를 포함하거나 이로 이루어진 TAP의 서열에서, T-세포 수용체와의 상호작용에 실질적으로 기여하지 않는 아미노산 잔기는, 다른 아미노산인 그의 도입이 T 세포 반응성에 실질적으로 영향을 주지 않고, 관련된 MHC에의 결합을 제거하지 않는 것인 다른 아미노산으로의 대체에 의해 변형될 수 있다.

[0100] TAP는 또한 분해를 방지하고, 안정성, 친화도 및/또는 흡수를 증가시키기 위하여 N- 및/또는 C-말단 캐핑될 수 있거나 변형될 수 있다. 따라서, 다른 양상에서, 본 개시내용은 화학식 Z^1-X-Z^2의 변형된 TAP을 제공하고, 여기서 X는 서열번호 1 내지 190, 바람직하게는 서열번호 97 내지 154의 아미노산 서열 중 하나를 포함하거나 이로 이루어진 TAP이다.

[0101] 실시형태에서, TAP의 아미노 말단 잔기(즉, N-말단 단부의 유리 아미노기)는, 예를 들어, 모이어티/화학기(Z^1)의 공유 부착에 의해 변형된다. Z^1은 1 내지 8개의 탄소의 직쇄 또는 분지형 알킬기 또는 아실기(R-CO-)일 수 있으며, 여기서 R은 소수성 모이어티(예컨대, 아세틸, 프로피오닐, 부탄일, 아이소-프로피오닐, 또는 아이소-부탄일), 또는 아로일기(Ar-CO-)(여기서 Ar은 아릴기임)이다. 실시형태에서, 아실기는 C_1-C_{16} 또는 C_3-C_{16} 아실기(선형 또는 분지형, 포화 또는 불포화), 추가의 실시형태에서, 포화된 C_1-C_6 아실기(선형 또는 분지형) 또는 불포화 C_3-C_6 아실기(선형 또는 분지형), 예를 들어, 아세틸기(CH_3-CO-, Ac)이다. 실시형태에서, Z^1은 존재하지 않는다. TAP의 카복시 말단 잔기(즉, TAP의 C-말단 단부에서의 유리 카복시기)는, 예를 들어, 아마이드화(OH기의 NH_2기로의 대체)에 의해 (예컨대, 분해에 대한 보호를 위하여) 변형될 수 있고, 따라서 이러한 경우에 Z^2는 NH_2기이다. 실시형태에서, Z^2는 하이드록사메이트기, 나이트릴기, 아마이드(1차, 2차 또는 3차)기, 1 내지 10개의 탄소의 지방족 아민, 예컨대, 메틸 아민, 아이소-부틸아민, 아이소-발레릴아민 또는 사이클로헥실아민, 방향족 또는 아릴알킬 아민, 예컨대, 아닐린, 나프틸아민, 벤질아민, 신나밀아민, 또는 페닐에틸아민, 알코올 또는 CH_2OH일 수 있다. 실시형태에서, Z^2는 존재하지 않는다. 실시형태에서, TAP는 서열번호 1 내지 190, 바람직하게는 서열번호 97 내지 154의 아미노산 서열 중 하나를 포함한다. 실시형태에서, TAP는 서열번호 1 내지 190, 바

람직하게는 서열번호 97 내지 154의 아미노산 서열 중 하나로 이루어지고, 여기서 Z^1 및 Z^2는 존재하지 않는다.

[0102] 다른 양상에서, 본 개시내용은, HLA-A*01:01 분자에 결합하고, 서열번호 48, 67, 89, 134, 151 또는 164, 서열번호 134 또는 151의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0103] 다른 양상에서, 본 개시내용은, HLA-A*02:01 분자에 결합하고, 서열번호 7, 11, 27, 32, 33, 34, 35, 4351, 52, 53, 54, 61, 65, 72, 77, 82, 86, 104, 108, 119, 123, 130, 132, 146, 150, 167, 168, 169, 171, 183 또는 188, 바람직하게는 서열번호 104, 108, 119, 123, 130, 132, 146 또는 150의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성(promiscuity)을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*02:05, HLA-A*02:06 및/또는 HLA-A*02:07 분자에 더 결합할 수 있다.

[0104] 다른 양상에서, 본 개시내용은, HLA-A*03:01 분자에 결합하고, 서열번호 5, 18, 19, 20, 42, 44, 74, 91, 105, 113, 126, 131, 159, 160, 180 또는 189, 바람직하게는 서열번호 105, 113, 126 또는 131의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*11:01 분자에 추가로 결합할 수 있다.

[0105] 다른 양상에서, 본 개시내용은, HLA-A*11:01 분자에 결합하고, 서열번호 6, 45, 96, 106, 121, 149 또는 152, 바람직하게는 서열번호 106, 121, 149 또는 152의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*03:01, HLA-A*31:01 및/또는 HLA-A*68:01 분자에 추가로 결합할 수 있다.

[0106] 다른 양상에서, 본 개시내용은, HLA-A*24:02 분자에 결합하고, 서열번호 13, 36, 71, 92, 95 또는 145, 바람직하게는 서열번호 145의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*23:01 분자에 추가로 결합할 수 있다.

[0107] 다른 양상에서, 본 개시내용은, HLA-A*26:01 분자에 결합하고, 서열번호 59 또는 서열번호 185의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*25:01 및/또는 HLA-A*66:01 분자에 추가로 결합할 수 있다.

[0108] 다른 양상에서, 본 개시내용은, HLA-A*29:02 분자에 결합하고, 서열번호 88, 99 또는 138, 바람직하게는 서열번호 99 또는 138의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*30:02 및/또는 HLA-B*15:02 분자에 추가로 결합할 수 있다.

[0109] 다른 양상에서, 본 개시내용은, HLA-A*30:01 분자에 결합하고, 서열번호 26, 29 또는 165의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0110] 다른 양상에서, 본 개시내용은, HLA-A*68:02 분자에 결합하고, 서열번호 50 또는 서열번호 148, 바람직하게는 서열번호 148의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0111] 다른 양상에서, 본 개시내용은, HLA-B*07:02 분자에 결합하고, 서열번호 8, 25, 40, 55, 56, 73, 78, 79, 80, 81, 98, 100, 107, 110, 111, 118, 120, 128, 129, 157, 161, 163, 179 또는 184, 바람직하게는 서열번호 98, 100, 107, 110, 111, 118, 120, 128 또는 129의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*35:02, HLA-B*35:03, HLA-B*55:01 및/또는 HLA-B*56:01 분자에 추가로 결합할 수 있다.

[0112] 다른 양상에서, 본 개시내용은, HLA-B*08:01 분자에 결합하고, 서열번호 4, 14, 15, 17, 서열번호 21, 22, 23, 28, 31, 47, 49, 101, 103, 114, 137, 139, 147, 156, 170, 181 또는 182, 바람직하게는 서열번호 101, 103,

114, 137, 139 또는 147의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0113] 다른 양상에서, 본 개시내용은, HLA-B*14:01 분자에 결합하고, 서열번호 2, 39, 58, 66, 124, 133, 136, 141, 162, 175, 176 또는 177, 바람직하게는 서열번호 124, 133, 136 또는 141의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0114] 다른 양상에서, 본 개시내용은, HLA-B*15:01 분자에 결합하고, 서열번호 10, 57, 62 또는 94의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*15:02, HLA-B*15:03 및/또는 HLA-B*46:01 분자에 추가로 결합할 수 있다.

[0115] 다른 양상에서, 본 개시내용은, HLA-B*27:05 분자에 결합하고, 서열번호 1 또는 144, 바람직하게는 서열번호 144의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*27:02 분자에 추가로 결합할 수 있다.

[0116] 다른 양상에서, 본 개시내용은, HLA-B*38:01 분자에 결합하고, 서열번호 4, 64 또는 84의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*39:01 분자에 추가로 결합할 수 있다.

[0117] 다른 양상에서, 본 개시내용은, HLA-B*40:01 분자에 결합하고, 서열번호 75 또는 76의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*18:01, HLA-B*40:02, HLA-B*41:02, HLA-B*44:02, HLA-B*44:03 및/또는 HLA-B*45:01 분자에 추가로 결합할 수 있다.

[0118] 다른 양상에서, 본 개시내용은, HLA-B*44:03 분자에 결합하고, 서열번호 127의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*18:01, HLA-B*40:01, HLA-B*40:02, HLA-B*41:02, HLA-B*44:02 및/또는 HLA-B*45:01 분자에 추가로 결합할 수 있다.

[0119] 다른 양상에서, 본 개시내용은, HLA-B*51:01 분자에 결합하고, 서열번호 12, 24, 38, 68, 115, 116 또는 190, 바람직하게는 서열번호 115 또는 116의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*35:02, HLA-B*35:03, HLA-B*52:01, HLA-B*53:01, HLA-B*55:01 및/또는 HLA-B*56:01 분자에 추가로 결합할 수 있다.

[0120] 다른 양상에서, 본 개시내용은, HLA-B*57:01 분자에 결합하고, 서열번호 3, 16, 41, 63, 69, 93, 122, 125, 135, 153 또는 183, 바람직하게는 122, 125, 135 또는 153의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-A*32:01 및/또는 HLA-B*58:01 분자에 추가로 결합할 수 있다.

[0121] 다른 양상에서, 본 개시내용은, HLA-B*57:03 분자에 결합하고, 서열번호 60 또는 172의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다.

[0122] 다른 양상에서, 본 개시내용은, HLA-C*03:03 분자에 결합하고, 서열번호 37, 112 또는 140, 바람직하게는 서열번호 112 또는 140의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*46:01, HLA-C*03:02, HLA-C*03:04, HLA-C*08:01, HLA-C*08:02, HLA-C*12:02, HLA-C*12:03, HLA-C*15:02 및/또는 HLA-C*16:01 분자에 추가로 결합할 수 있다.

[0123] 다른 양상에서, 본 개시내용은, HLA-C*05:01 분자에 결합하고, 서열번호 150의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성

을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-C*08:01 및/또는 HLA-C*08:02 분자에 추가로 결합할 수 있다.

[0124] 다른 양상에서, 본 개시내용은, HLA-C*06:02 분자에 결합하고, 서열번호 9, 70, 87, 142, 154 또는 166, 바람직하게는 서열번호 142 또는 154의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*27:02, HLA-C*07:01 및/또는 HLA-C*07:02 분자에 추가로 결합할 수 있다.

[0125] 다른 양상에서, 본 개시내용은, HLA-C*07:01 분자에 결합하고, 서열번호 142, 143, 155, 178 또는 186), 바람직하게는 서열번호 142 또는 143의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*27:02, HLA-C*07:01, HLA-C*07:02 및/또는 HLA-C*14:02 분자에 추가로 결합할 수 있다.

[0126] 다른 양상에서, 본 개시내용은, HLA-C*07:02 분자에 결합하고, 서열번호 30, 83, 85, 90, 109, 117 또는 158, 바람직하게는 서열번호 109 또는 117의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*27:02, HLA-C*07:01, HLA-C*07:02 및/또는 HLA-C*14:02 분자에 추가로 결합할 수 있다.

[0127] 다른 양상에서, 본 개시내용은, HLA-C*08:02 분자에 결합하고, 서열번호 97, 102 또는 174, 바람직하게는 서열번호 97 또는 102의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-C*03:03, HLA-C*03:04, HLA-C*05:01, HLA-C*08:01 및/또는 HLA-C*15:02 분자에 추가로 결합할 수 있다.

[0128] 다른 양상에서, 본 개시내용은, HLA-C*12:03 분자에 결합하고, 서열번호 173의 서열을 포함하거나 또는 이로 이루어진, 백혈병 TAP(또는 종양-특이적 펩타이드), 바람직하게는 AML TAP를 제공한다. HLA 대립유전자가 난잡성을 나타내기 때문에(소정의 HLA 대립유전자는 유사한 에피토프를 제공함, **표 4** 참조), 위에서 확인된 TAP는 HLA-B*46:01, HLA-C*03:02, HLA-C*03:03, HLA-C*03:04, HLA-C*08:01, HLA-C*12:03, HLA-C*15:02 및/또는 HLA-C*16:01 분자에 추가로 결합할 수 있다.

[0129] 실시형태에서, TAP는 비번역 전사된 영역(UTR), 즉 3'-UTR 또는 5'-UTR 영역에 위치된 서열에 의해 인코딩된다. 다른 실시형태에서, TAP는 인트론에 위치된 서열에 의해 인코딩된다. 다른 실시형태에서, TAP는 유전자간 영역에 위치된 서열에 의해 인코딩된다. 다른 실시형태에서, TAP는 엑손에 위치된 서열에 의해 인코딩되고 프레임시프트로부터 기원한다.

[0130] 본 개시내용의 TAP는 TAP를 인코딩하는 핵산을 포함하는 숙주 세포에서의 발현(재조합 발현)에 의해, 또는 화학적 합성(예컨대, 고체상 펩타이드 합성)에 의해 제조될 수 있다. 펩타이드는 당업계에 널리 공지된 수동 및/또는 자동 고체상 절차에 의해 용이하게 합성될 수 있다. 적합한 합성법은, 예를 들어, "T-boc" 또는 "Fmoc" 절차를 사용함으로써 수행될 수 있다. 고체상 합성을 위한 기술 및 절차는, 예를 들어, 문헌[Solid Phase Peptide Synthesis: A Practical Approach, by E. Atherton and R. C. Sheppard, published by IRL, Oxford University Press, 1989]에 기재되어 있다. 대안적으로, MiHA 펩타이드는, 예를 들어, 문헌[Liu *et al.*, *Tetrahedron Lett*. 37: 933-936, 1996; Baca *et al.*, *J. Am. Chem. Soc*. 117: 1881-1887, 1995; Tam *et al.*, *Int. J. Peptide Protein Res*. 45: 209-216, 1995; Schnolzer and Kent, *Science* 256: 221-225, 1992; Liu and Tam, *J. Am. Chem. Soc*. 116: 4149-4153, 1994; Liu and Tam, *Proc. Natl. Acad. Sci. USA* 91: 6584-6588, 1994; 및 Yamashiro and Li, *Int. J. Peptide Protein Res*. 31: 322-334, 1988)]에 기재된 바와 같이, 세그먼트 축합에 의해 제조될 수 있다. TAP를 합성하는 데 유용한 다른 방법은 문헌[Nakagawa *et al.*, *J. Am. Chem. Soc*. 107: 7087-7092, 1985]에 기재되어 있다. 실시형태에서, TAP은 화학적으로 합성된다(합성 펩타이드). 본 개시내용의 또 다른 실시형태는 본 명세서에서 정의된 아미노산 서열로 이루어지거나, 또는 본질적으로 이로 이루어지고, 약제학적으로 허용 가능한 염으로서 합성에 의해 제조된(예컨대, 합성된) 비천연형 펩타이드에 관한 것이다. 본 개시내용에 따른 TAP의 염은 생체내에서의 상태(들)의 펩타이드와는 실질적으로 상이한데, 그 이유는 생체내에서 생성된 것과 같은 펩타이드는 염이 없기 때문이다. 펩타이드의 비천연 염 형태는, 특히, 펩타이드를 포함하는 약제학적 조성물, 예컨대, 본 명세서에 개시된 바와 같은 펩타이드 백신의

맥락에서, 펩타이드의 가용성을 조정할 수 있다. 바람직하게는, 염은 펩타이드의 약제학적으로 허용 가능한 염이다.

[0131] 실시형태에서, 본 명세서에서 언급된 TAP는 실질적으로 순수한 것이다. 화합물은 자연적으로 이를 수반하는 성분으로부터 분리되었을 때 "실질적으로 순수한" 것이다. 전형적으로, 화합물은, 샘플 중 전체 물질의 적어도 60 중량%, 더 일반적으로, 75 중량%, 80 중량% 또는 85 중량%, 바람직하게는, 90 중량% 초과, 더 바람직하게는, 95 중량% 초과일 때, 실질적으로 순수한 것이다. 따라서, 예를 들어, 화학적으로 합성되거나, 또는 재조합 기술에 의해 제조된 폴리펩타이드는 일반적으로 그의 자연적으로 회합된 성분, 예컨대, 그의 공급원 거대분자의 성분을 실질적으로는 함유하지 않을 것이다. 핵산 분자는, 통상적으로 핵산이 유래되는 유기체의 천연형 게놈에서 인접해 있는 코딩 서열과 바로 인접해 있지 않을 때(즉, 공유적으로 연결되어 있지 않을 때), 실질적으로 순수한 것이다. 실질적으로 순수한 화합물은, 예를 들어, 천연 공급원으로부터의 추출에 의해; 펩타이드 화합물을 인코딩하는 재조합 핵산 분자의 발현에 의해; 또는 화학적 합성에 의해 얻어질 수 있다. 순도는 임의의 적절한 방법, 예컨대, 칼럼 크로마토그래피, 겔 전기영동, HPLC 등을 사용하여 측정될 수 있다. 실시형태에서, TAP는 용액 중에 있다. 다른 실시형태에서, TAP는 고체 형태, 예컨대, 동결건조된 것이다.

[0132] 다른 양상에서, 본 개시내용은 본 명세서에서 언급된 TAP 또는 종양 항원 전구체-펩타이드를 인코딩하는 (단리된) 핵산을 더 제공한다. 실시형태에서, 핵산은 약 21개의 뉴클레오타이드 내지 약 45개의 뉴클레오타이드, 약 24 내지 약 45개의 뉴클레오타이드, 예를 들어 24, 27, 30, 33, 36, 39, 42 또는 45개의 뉴클레오타이드를 포함한다. 본 명세서에서 사용되는 바와 같은 "단리된"은, 펩타이드 또는 핵산 분자가 당해 분자, 또는 천연형 공급원 거대분자(예컨대, 다른 핵산, 단백질, 지질, 당 등 포함)의 자연 환경에 존재하는 다른 성분으로부터 분리된 펩타이드 또는 핵산 분자를 지칭한다. 본 명세서에서 사용되는 바와 같은 "합성"은, 펩타이드 또는 핵산 분자가 그의 자연 공급원으로부터 단리된 것이 아니라, 예컨대, 재조합 기술을 통해 또는 화학적 합성을 이용하여 제조된 것을 지칭한다. 본 개시내용의 핵산은 본 개시내용의 TAP의 재조합 발현을 위해 사용될 수 있고, 숙주 세포 내로 형질감염될 수 있는 벡터 또는 플라스미드, 예컨대, 클로닝 벡터 또는 발현 벡터에 포함될 수 있다. 실시형태에서, 본 개시내용은 본 개시내용의 TAP를 인코딩하는 핵산 서열을 포함하는 클로닝, 발현 또는 바이러스 벡터 또는 플라스미드를 제공한다. 대안적으로, 본 개시내용의 TAP를 인코딩하는 핵산은 숙주 세포의 게놈 내로 도입될 수 있다. 어느 경우든지, 숙주 세포는 핵산에 의해 인코딩된 TAP 또는 단백질을 발현한다. 본 명세서에서 사용되는 바와 같은 용어 "숙주 세포"는 특정 대상체 세포뿐만 아니라, 이러한 세포의 자손 또는 잠재적 자손을 지칭한다. 숙주 세포는 본 명세서에 기재된 TAP를 발현할 수 있는 임의의 원핵 세포(예컨대, 이. 콜라이(*E. coli*)) 또는 진핵 세포(예컨대, 곤충 세포, 효모 또는 포유동물 세포)일 수 있다. 벡터 또는 플라스미드는 삽입된 코딩 서열의 전사 및 번역에 필요한 요소를 함유하고, 다른 성분, 예컨대, 저항 유전자, 클로닝 부위 등을 함유할 수 있다. 당업자에게 널리 공지된 방법을 사용하여, 펩타이드 또는 폴리펩타이드를 인코딩하는 서열, 및 이에 작동 가능하게 연결된 적절한 전사 및 번역 제어/조절 요소를 함유하는 발현 벡터를 작제할 수 있다. 이러한 방법은 시험관내 재조합 DNA 기술, 합성 기술, 및 시험관내 유전자 재조합을 포함한다. 이러한 기술은 문헌[Sambrook. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, N.Y.]에 기재되어 있다. "작동 가능하게 연결된"은, 성분들의 정상적인 기능이 수행될 수 있도록 하는, 성분, 특히 뉴클레오타이드 서열의 병치를 지칭한다. 따라서, 조절 서열에 작동 가능하게 연결된 코딩 서열은, 당해 코딩 서열이 조절 서열의 조절 제어, 즉, 전사 및/또는 번역 제어하에 발현될 수 있는 뉴클레오타이드 서열의 배치를 지칭한다. 본 명세서에서 사용되는 바와 같은 "조절/제어 영역" 또는 "조절/제어 서열"은, 코딩 핵산의 발현의 조절에 관여하는 비-코딩 뉴클레오타이드 서열을 지칭한다. 따라서, 용어 조절 영역은 프로모터 서열, 조절 단백질 결합 부위, 상류 활성제 서열 등을 포함한다. 벡터(예컨대, 발현 벡터)는 각각의 숙주 세포에서 효과적인 유전자 전사 및 번역을 위한 필요한 5' 상류 및 3' 하류 조절 요소, 예컨대, 프로모터 서열, 예컨대, CMV, PGK 및 EFla 프로모터, 리보솜 인식 및 결합 TATA 박스, 및 3' UTR AAUAAA 전사 종결 서열을 가질 수 있다. 다른 적합한 프로모터는 심미안 바이러스 40(SV40) 조기 프로모터, 마우스 유방 종양 바이러스(MMTV), HIV LTR 프로모터, MoMuLV 프로모터, 조류 백혈병 바이러스 프로모터, EBV 극초기 프로모터 및 라우스 육종 바이러스 프로모터의 구성적 프로모터를 포함한다. 액틴 프로모터, 미오신 프로모터, 헤모글로빈 프로모터 및 크레아틴 키나제 프로모터를 포함하지만, 이들로 제한되지 않는 인간 유전자 프로모터가 또한 사용될 수 있다. 소정의 실시형태에서 유도성 프로모터는 TAP를 발현하는 벡터의 일부로서 상정된다. 이것은 관심 폴리뉴클레오타이드 서열의 발현을 켜거나 발현을 끄는 분자 스위치를 제공한다. 유도성 프로모터의 예는 메탈로티오닌 프로모터, 글루코코르티코이드 프로모터, 프로게스테론 프로모터 또는 테트라사이클린 프로모터를 포함하지만, 이들로 제한되지 않는다. 벡터의 예는 플라스미드, 자율 복제 서열 및 전이성 요소(transposable element)이다.

추가의 예시적인 벡터는, 제한 없이, 플라스미드, 파지미드, 코스미드, 인공 염색체, 예컨대, 효모 인공 염색체(YAC), 박테리아 인공 염색체(BAC), 또는 P1-유도 인공 염색체(PAC), 박테리오파지, 예컨대, 람다 파지 또는 M13 파지, 및 동물 바이러스를 포함한다. 벡터로서 유용한 동물 바이러스의 범주의 예는, 제한 없이, 레트로바이러스(렌티바이러스 포함), 아데노바이러스, 아데노-연관 바이러스, 헤르페스바이러스(예컨대, 단순포진 바이러스), 폭스바이러스, 바큘로바이러스, 유두종바이러스 및 파포바바이러스(예컨대, SV40)를 포함한다. 발현 벡터의 예는 포유동물 세포에서 발현을 위한 Lenti-X™ Bicistronic Expression System (Neo) 벡터(Clontrch), pClneo 벡터(Promega); 포유동물 세포에서 렌티바이러스-매개 유전자 수송 및 발현을 위한 pLenti4/V5-DEST™, pLenti6/V5-DEST™ 및 pLenti6.2N5-GW/lacZ(Invitrogen)이다. 본 명세서에 개시된 TAP의 코딩 서열은 포유동물 세포에서 TAP의 발현을 위하여 이러한 발현 벡터에 결찰될 수 있다.

[0133] 소정의 실시형태에서, 본 개시내용의 TAP를 인코딩하는 핵산은 바이러스 벡터에 제공된다. 바이러스 벡터는 레트로바이러스, 렌티바이러스, 또는 거품형성 바이러스(foamy virus)로부터 유래된 것들일 수 있다. 본 명세서에서 사용되는 바와 같이, 용어 "바이러스 벡터"는 바이러스 기원의 적어도 하나의 요소를 포함하고 바이러스 벡터 입자 내로 패키징되는 능력을 갖는 핵산 벡터 작제물을 지칭한다. 바이러스 벡터는 비필수 바이러스 유전자 대신에 본 명세서에 기재된 다양한 단백질에 대한 코딩 서열을 함유할 수 있다. 벡터 및/또는 입자는 DNA, RNA 또는 다른 핵산을 시험관내 또는 생체내에서 세포에 수송할 목적으로 사용될 수 있다. 수많은 형태의 바이러스 벡터가 당업계에 공지되어 있다.

[0134] 실시형태에서, 본 개시내용의 TAP를 인코딩하는 핵산(DNA, RNA)은 리포솜 또는 임의의 다른 적합한 비히클 내에 포함된다.

[0135] 다른 양상에서, 본 개시내용은 서열번호 1 내지 190, 바람직하게는 서열번호 97 내지 154의 TAP의 하나 이상을 포함하는(즉, 제시하는 또는 이에 결합된) MHC 클래스 I 분자를 제공한다. 실시형태에서, MHC 클래스 I 분자는 HLA-A1 분자, 추가의 실시형태에서 HLA-A*01:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A2 분자, 추가의 실시형태에서 HLA-A*02:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A3 분자, 추가의 실시형태에서 HLA-A*03:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A11 분자, 추가의 실시형태에서 HLA-A*11:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A24 분자, 추가의 실시형태에서 HLA-A*24:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A26 분자, 추가의 실시형태에서 HLA-A*26:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A29 분자, 추가의 실시형태에서 HLA-A*29:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A30 분자, 추가의 실시형태에서 HLA-A*30:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-A68 분자, 추가의 실시형태에서 HLA-A*68:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B07 분자, 추가의 실시형태에서 HLA-B*07:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B08 분자, 추가의 실시형태에서 HLA-B*08:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B14 분자, 추가의 실시형태에서 HLA-B*14:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B15 분자, 추가의 실시형태에서 HLA-B*15:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B27 분자, 추가의 실시형태에서 HLA-B*27:05 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B38 분자, 추가의 실시형태에서 HLA-B*38:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B40 분자, 추가의 실시형태에서 HLA-B*40:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B44 분자, 추가의 실시형태에서 HLA-B*44:02 또는 HLA-B*44:03 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-B57 분자, 추가의 실시형태에서 HLA-B*57:01 또는 HLA-B*57:03 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C03 분자, 추가의 실시형태에서 HLA-C*03:03 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C04 분자, 추가의 실시형태에서 HLA-C*04:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C05 분자, 추가의 실시형태에서 HLA-C*05:01 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C06 분자, 추가의 실시형태에서 HLA-C*06:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C07 분자, 추가의 실시형태에서 HLA-C*07:01 또는 HLA-C*07:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C08 분자, 추가의 실시형태에서 HLA-C*08:02 분자이다. 다른 실시형태에서, MHC 클래스 I 분자는 HLA-C12 분자, 추가의 실시형태에서 HLA-C*12:03 분자이다.

[0136] 실시형태에서, TAP는 MHC 클래스 I 분자에 비공유적으로 결합되어 있다(즉, TAP는 MHC 클래스 I 분자의 펩타이드 결합 홈/포켓 내로 로딩되어 있거나, 또는 이에 비공유적으로 결합되어 있다). 다른 실시형태에서, TAP는 MHC 클래스 I 분자(알파 사슬)에 공유적으로 부착/결합되어 있다. 이러한 작제물에서, TAP 및 MHC 클래스 I 분자(알파 사슬)는, 전형적으로 짧은(예컨대, 5 내지 20개, 바람직하게는 약 8 내지 12, 예컨대, 10개의 잔기의) 가요성 링커 또는 스페이서(예컨대, 폴리글리신 링커)와 함께 합성 융합 단백질로서 생성된다. 다른 양상에서, 본 개시내용은 MHC 클래스 I 분자(알파 사슬)에 융합된 본 명세서에서 정의된 TAP를 포함하는 융합 단백질을 인

코딩하는 핵산을 제공한다. 실시형태에서, MHC 클래스 I 분자(알파 사슬) - 펩타이드 복합체는 다량체화된다. 따라서, 다른 양상에서, 본 개시내용은 본 명세서에서 언급된 TAP가 (공유적으로 또는 공유적이 아니게) 로딩된 MHC 클래스 I 분자의 다량체를 제공한다. 이러한 다량체는 다량체의 검출을 허용하는 태그, 예를 들어, 형광성 태그에 부착될 수 있다. MHC 이량체, 사량체, 오량체, 팔량체 등을 비롯한 MHC 다량체의 생성을 위한 다수의 전략법이 개발되었다(문헌[Bakker and Schumacher, *Current Opinion in Immunology* 2005, 17:428-433]에서 검토됨). MHC 다량체는, 예를 들어, 항원-특이적 T 세포의 검출 및 정제에 유용하다. 따라서, 다른 양상에서, 본 개시내용은, 본 명세서에 정의된 TAP에 특이적인 $CD8^+$ T 림프구를 검출 또는 정제(단리, 농축)하는 방법을 제공하되, 해당 방법은, 세포 집단을, TAP가 (공유적으로 또는 공유적이 아니게) 로딩된 MHC 클래스 I 분자의 다량체와 접촉시키는 단계; 및 MHC 클래스 I 다량체에 의해 결합된 $CD8^+$ T 림프구를 검출 또는 단리하는 단계를 포함한다. MHC 클래스 I 다량체에 의해 결합된 $CD8^+$ T 림프구는 공지된 방법, 예를 들어, 형광 활성화 세포 분류(fluorescence activated cell sorting: FACS) 또는 자기 활성화 세포 분류(magnetic activated cell sorting: MACS)를 이용하여 단리될 수 있다.

[0137] 또 다른 양상에서, 본 개시내용은, 본 개시내용의 본 명세서에서 언급된 핵산, 벡터 또는 플라스미드, 즉, 하나 이상의 TAP를 인코딩하는 핵산 또는 벡터를 포함하는, 세포(예컨대, 숙주 세포), 실시형태에서, 단리된 세포를 제공한다. 다른 양상에서, 본 개시내용은, 그의 표면에서, 본 개시내용에 따른 TAP에 결합 또는 이를 제시하는 MHC 클래스 I 분자(예컨대, 위에서 개시된 대립유전자 중 하나의 MHC 클래스 I 분자)를 발현하는 세포를 제공한다. 일 실시형태에서, 숙주 세포는 진핵 세포, 예컨대, 포유동물 세포, 바람직하게는, 인간 세포, 세포주 또는 불멸화 세포이다. 다른 실시형태에서, 세포는 항원-제시 세포(APC)이다. 일 실시형태에서, 숙주 세포는 1차 세포, 세포주 또는 불멸화 세포이다. 다른 실시형태에서, 세포는 항원-제시 세포(APC)이다. 핵산 및 벡터는 통상의 형질전환 또는 형질감염 기술을 통해서 세포에 도입될 수 있다. 용어 "형질전환" 및 "형질감염"은 인산칼슘 또는 염화칼슘 공침, DEAE-덱스트란 매개 형질감염, 리포펙션, 전기천공, 미세주입 및 바이러스 매개 형질감염을 비롯한, 외래 핵산을 숙주 세포 내로 도입하기 위한 기술을 지칭한다. 숙주 세포를 형질전환 또는 형질감염시키는 데 적합한 방법은, 예를 들어, 문헌[Sambrook *et al.* (상기 참조)] 및 다른 실험실용 매뉴얼에서 찾을 수 있다. 핵산을 생체내에서 포유동물 세포 내로 도입하기 위한 방법이 또한 공지되어 있고, 유전자 요법을 위해 본 개시내용의 벡터 또는 플라스미드를 대상체에게 전달하는 데 사용될 수 있다.

[0138] APC와 같은 세포는 당업계에 공지된 다양한 방법을 이용하여 하나 이상의 TAP로 로딩될 수 있다. 본 명세서에서 사용되는 바와 같은 TAP를 "세포에 로딩하는" 것은, TAP를 인코딩하는 RNA 또는 DNA, 또는 TAP를 세포 내로 형질감염시키거나, 또는 대안적으로 TAP를 인코딩하는 핵산으로 APC를 형질전환시키는 것을 의미한다. 세포는 또한 세포 표면에 제시된 MHC 클래스 I 분자에 직접 결합할 수 있는 외인성 TAP와 세포(예컨대, 펩타이드-펄싱된 세포)를 접촉시킴으로써 로딩될 수 있다. TAP는 또한 MHC 클래스 I 분자에 의한 그의 제시를 촉진시키는 도메인 또는 모티프에, 예를 들어, 소포체(ER) 복구 신호, C-말단 Lys-Asp-Glu-Leu 서열에 융합될 수 있다(문헌[Wang *et al.*, *Eur J Immunol.* 2004 Dec;34(12):3582-94] 참조).

[0139] 다른 양상에서, 본 개시내용은, 본 명세서에 정의된 TAP 중 어느 하나 또는 이들의 임의의 조합(또는 상기 펩타이드(들)를 인코딩하는 핵산))을 포함하는 조성물 또는 펩타이드 조합물/풀을 제공한다. 실시형태에서, 조성물은 본 명세서에 정의된 TAP의 임의의 조합물(2, 3, 4, 5, 6, 7, 8, 9, 10개 이상의 TAP의 임의의 조합물), 또는 상기 TAP를 인코딩하는 핵산의 조합물을 포함한다. 본 명세서에 정의된 TAP의 임의의 조합/하위 조합을 포함하는 조성물은 본 개시내용에 의해 포괄된다. 다른 실시형태에서, 조합물 또는 풀은 하나 이상의 공지된 종양 항원을 포함할 수 있다.

[0140] 따라서, 다른 양상에서, 본 개시내용은, 본 명세서에 정의된 TAP 중 어느 하나 또는 이들의 임의의 조합물 및 MHC 클래스 I 분자(예컨대, 위에서 개시된 대립유전자 중 하나의 MHC 클래스 I 분자)를 발현하는 세포를 포함하는 조성물을 제공한다. 본 개시내용에서 사용하기 위한 APC는 특정 유형의 세포로 제한되지 않고, $CD8^+$ T 림프구에 의해 인식되도록 하기 위해 그의 세포 표면 상에 단백질성 항원을 제시하는 것으로 공지된, 전문 APC, 예컨대, 수지상 세포(DC), 랑게르한스 세포, 대식세포 및 B 세포를 포함한다. 예를 들어, APC는 말초 혈액 단핵구로부터 DC를 유도한 후, 시험관내, 생체외 또는 생체내에서 TAP를 접촉(자극)시킴으로써 얻어질 수 있다. APC는 또한 본 개시내용의 TAP 중 하나 이상이 대상체에게 투여되고, TAP를 제시하는 APC가 대상체의 체내에서 유도되는 경우 생체내에서 TAP를 제시하도록 활성화될 수 있다. 어구 "APC를 유도하는" 또는 "APC를 자극시키는"은 세포를 하나 이상의 TAP, 또는 TAP를 인코딩하는 핵산과 접촉시키거나, 또는 그를 세포에 로딩하여 TAP가 MHC 클래스 I 분자에 의해 그의 표면에서 제시되도록 하는 것을 포함한다. 본 명세서에서 언급된 바와 같이, 본 개시

내용에 따르면, TAP는, 예를 들어, TAP(천연 단백질 포함)의 서열을 포함하는 더 긴 펩타이드/폴리펩타이드를 이용하여 간접적으로 로딩될 수 있고, 이어서, APC 내부에서 (예컨대, 프로테아제에 의해) 프로세싱되어 세포 표면에서 TAP/MHC 클래스 I 복합체를 생성한다. APC에 TAP를 로딩하고, APC가 TAP를 제시하게 허용한 후, APC를 백신으로서 대상체에게 투여할 수 있다. 예를 들어, 생체외 투여는 (a) 제1 대상체로부터 APC를 수집하는 단계, (b) 단계 (a)의 APC를 TAP와 접촉/로딩하여 APC의 표면에서 MHC 클래스 I/TAP 복합체를 형성하는 단계; 및 (c) 치료를 필요로 하는 제2 대상체에게 펩타이드-로딩된 APC를 투여하는 단계를 포함할 수 있다.

[0141] 제1 대상체 및 제2 대상체는 동일한 대상체(예컨대, 자가 백신)일 수 있거나, 또는 상이한 대상체(예컨대, 동종 백신)일 수 있다. 대안적으로, 본 개시내용에 따르면, 항원-제시 세포를 유도하기 위한 조성물(예컨대, 약제학적 조성물)을 제조하기 위한 본 명세서에 기재된 TAP(또는 이들의 조합물)의 용도가 제공된다. 또한, 본 개시내용은 항원-제시 세포를 유도하기 위한 약제학적 조성물을 제조하는 방법 또는 공정을 제공하되, 해당 방법 또는 공정은 TAP, 또는 이들의 조합물을 약제학적으로 허용 가능한 담체와 혼합 또는 제형화하는 단계를 포함한다. 본 명세서에 정의된 TAP 중 어느 하나, 또는 이들의 임의의 조합물이 로딩된, MHC 클래스 I 분자(예컨대, HLA-A1, HLA-A2, HLA-A3, HLA-A11, HLA-A24, HLA-A25, HLA-A29, HLA-A32, HLA-B07, HLA-B08, HLA-B14, HLA-B15, HLA-B18, HLA-B39, HLA-B40, HLA-B44, HLA-C03, HLA-C04, HLA-C05, HLA-C06, HLA-C07, HLA-C12 또는 HLA-C14 분자)를 발현하는 APC와 같은 세포는 $CD8^+$ T 림프구, 예를 들어, 자가 $CD8^+$ T 림프구를 자극/증폭시키는 데 사용될 수 있다. 따라서, 다른 양상에서, 본 개시내용은 본 명세서에 정의된 TAP 중 어느 하나, 또는 이들의 임의의 조합물(또는 핵산 또는 이를 인코딩하는 벡터); MHC 클래스 I 분자 및 T 림프구, 더욱 구체적으로, $CD8^+$ T 림프구를 발현하는 세포(예컨대, $CD8^+$ T 림프구를 포함하는 세포 집단)를 포함하는 조성물을 제공한다.

[0142] 실시형태에서, 조성물은 완충제, 부형제, 담체, 희석제 및/또는 배지(예컨대, 배양 배지)를 더 포함한다. 추가의 실시형태에서, 완충제, 부형제, 담체, 희석제 및/또는 배지는 약제학적으로 허용 가능한 완충제(들), 부형제(들), 담체(들), 희석제(들) 및/또는 배지(들)이다. 본 명세서에서 사용되는 바와 같은 "약제학적으로 허용 가능한 완충제, 부형제, 담체, 희석제 및/또는배지"는 생리적으로 상용성이고, 활성 성분(들)의 생물학적 활성의 효과를 방해하지 않으며, 대상체에게 비독성인, 임의의 모든 용매, 완충제, 결합제, 윤활제, 충전제, 증점제, 붕해제, 가소화제, 코팅제, 장벽 층 제형, 윤활제, 안정화제, 방출-지연제, 분산 매질, 코팅제, 항박테리아제 및 항진균제, 등장제 등을 포함한다. 약제학적 활성 물질을 위한 이러한 배지 및 제제의 사용은 당업계에 널리 공지되어 있다(Rowe et al., Handbook of pharmaceutical excipients, 2003, 4^{th} edition, Pharmaceutical Press, London UK). 임의의 통상의 배지 또는 제제가 활성 화합물(펩타이드, 세포)과 비상용성이지 않는 한, 본 개시내용의 조성물에서의 이의 용도가 고려된다. 실시형태에서, 완충제, 부형제, 담체 및/또는 배지는 비천연형 완충제, 부형제, 담체 및/또는 배지이다. 실시형태에서, 본 명세서에 정의된 TAP 중 하나 이상, 또는 상기 하나 이상의 TAP를 인코딩하는 핵산(예컨대, mRNA)은, 리포솜, 예컨대, 양이온성 리포솜 또는 적합한 다른 담체 내에 포함되거나, 또는 이에 복합체화된다(예컨대, 문헌[Vitor MT *et al.*, *Recent Pat Drug Deliv Formul*. 2013 Aug;7(2):99-110] 참조).

[0143] 다른 양상에서, 본 개시내용은 본 명세서에 정의된 TAP 중 어느 하나 중 하나 이상의 것, 또는 이들의 임의 조합물(또는 상기 펩타이드(들)를 인코딩하는 핵산), 및 완충제, 부형제, 담체, 희석제 및/또는 배지를 포함하는 조성물을 제공한다. 세포(예컨대, APC, T 림프구)를 포함하는 조성물의 경우, 조성물은 생존 세포의 유지를 허용하는 적합하는 배지를 포함한다. 이러한 배지의 대표적인 예는 염수 용액, 얼스 균형 잡힌 염 용액(Earl's Balanced Salt Solution)(Life Technologies®) 또는 PlasmaLyte®(Baxter International®)를 포함한다. 실시형태에서, 조성물(예컨대, 약제학적 조성물)은 "면역원성 조성물", "백신 조성물" 또는 "백신"이다. 본 명세서에서 사용되는 바와 같은 용어 "면역원성 조성물", "백신 조성물" 또는 "백신"은 하나 이상의 TAP 또는 백신 벡터를 포함하고 대상체에게 투여된 경우, 그에 존재하는 하나 이상의 TAP에 대하여 면역 반응을 유도할 수 있는 조성물 또는 제형을 지칭한다. 포유동물에서 면역 반응을 유도하기 위한 백신접종 방법은 백신 분야에 공지된 임의의 통상의 경로, 예컨대, 점막(예컨대, 눈, 비강내, 폐, 경구, 위, 장, 직장, 질 또는 요로) 표면을 통해, 비경구적(예컨대, 피하, 진피내, 근육내, 정맥내, 또는 복강내) 경로를 통해, 또는 국소 투여를 통해(예컨대, 경피 전달 시스템, 예컨대, 패치를 통해) 투여되는 백신 또는 백신 벡터의 사용을 포함한다. 실시형태에서, TAP(또는 이의 조합물)는 TAP(들)의 면역원성을 증가시키기 위해 운반 단백질(접합체 백신)에 접합된다. 따라서, 본 개시내용은 TAP(또는 이의 조합물), 또는 TAP를 인코딩하는 핵산 또는 이들의 조합물, 및 운반 단백질을 포함하는 조성물(접합체)을 제공한다. 예를 들어, TAP(들) 또는 핵산(들)은 톨-유사 수용체(Toll-like receptor: TLR) 리간드(예컨대, 문헌[Zom *et al.*, *Adv Immunol*. 2012, 114: 177-201] 참조) 또는 중합체/덴드

리머(예컨대, 문헌[Liu *et al.*, *Biomacromolecules*. 2013 Aug 12;14(8):2798-806] 참조)에 접합 또는 복합체화될 수 있다. 실시형태에서, 면역원성 조성물 또는 백신은 애주번트를 더 포함한다. "애주번트"는, 면역원성 제제, 예컨대, 항원(본 개시내용에 따른 TAP, 핵산 및/또는 세포)에 첨가된 경우, 혼합물에의 노출시 숙주에서 상기 제제에 대한 면역 반응을 비특이적으로 증강 또는 강화시키는 물질을 지칭한다. 현재 백신 분야에서 사용되는 애주번트의 예는 (1) 무기염(알루미늄염, 예컨대, 인산알루미늄 및 수산화알루미늄, 인산칼슘 겔), 스쿠알렌, (2) 오일-기반 애주번트, 예컨대, 오일 에멀션 및 계면활성제 기반 제형, 예컨대, MF59(미세유동화된 세제 안정화 수중유 에멀션), QS21(정제된 사포닌), AS02[SBAS2](수중유 에멀션 + MPL + QS-21), (3) 미립자 애주번트, 예컨대, 비로좀(virosome)(인플루엔자 헤마글루티닌을 포함하는 단일층 리포솜 비히클), AS04(MPL과의 [SBAS4] 알루미늄염), ISCOMS(사포닌 및 지질의 구조화된 복합체), 폴리락타이드 코-글리콜라이드(polylactide co-glycolide: PLG), (4) 미생물 유도체(천연 및 합성), 예컨대, 모노포스포릴 리피드 A(MPL), 데톡스(Detox)(MPL + 엠. 플레이(*M. Phlei*) 세포벽 골격), AGP[RC-529](합성 아실화된 단당류), DC_Chol(리포솜으로 자기 조직화할 수 있는 지질 면역자극제), OM-174(지질 A 유도체), CpG 모티프(면역자극성 CpG 모티프를 함유하는 합성 올리고뉴클레오타이드), 변형된 LT 및 CT(비독성 애주번트 효과를 제공하는 유전자 변형된 박테리아 독소), (5) 내인성 인간 면역자극제, 예컨대, hGM-CSF 또는 hIL-12(인코딩된 단백질 또는 플라스미드로서 투여될 수 있는 사이토카인), 이뮤답틴(Immudaptin)(C3d 탠덤 어레이) 및/또는 (6) 불활성 비히클, 예컨대, 금 입자 등을 포함한다.

[0144] 실시형태에서, TAP(들) 또는 이를 포함하는 조성물은 동결건조된 형태이다. 다른 실시형태에서, TAP(들) 또는 이를 포함하는 조성물은 액체 조성물이다. 추가의 실시형태에서, TAP(들)는 조성물 중 약 0.01 ㎍/㎖ 내지 약 100 ㎍/㎖의 농도로 존재한다. 추가의 실시형태에서, TAP(들)는 조성물 중 약 0.2 ㎍/㎖ 내지 약 50 ㎍/㎖, 약 0.5 ㎍/㎖ 내지 약 10, 20, 30, 40 또는 50 ㎍/㎖, 약 1 ㎍/㎖ 내지 약 10 ㎍/㎖, 또는 약 2 ㎍/㎖의 농도로 존재한다.

[0145] 본 명세서에서 언급된 바와 같이, 본 명세서에서 정의된 TAP 중 어느 하나, 또는 이들의 임의의 조합물이 로딩되거나, 또는 이에 결합된 MHC 클래스 I 분자를 발현하는 APC와 같은 세포는, 생체내에서 또는 생체외에서 $CD8^+$ T 림프구를 자극/증폭시키기 위하여 사용될 수 있다. 따라서, 다른 양상에서, 본 개시내용은 본 명세서에서 언급된 MHC 클래스 I 분자/TAP 복합체와 상호작용할 수 있거나 또는 이와 결합할 수 있는 T 세포 수용체(TCR) 분자, 및 이러한 TCR 분자를 인코딩하는 핵산 분자, 및 이러한 핵산 분자를 포함하는 벡터를 제공한다. 본 개시내용에 따른 TCR은, 바람직하게는, 생세포의 표면에서 시험관내 또는 생체내에서 MHC 클래스 I 분자에 로딩되거나 당해 분자에 의해 제시된 TAP와 특이적으로 상호작용하거나 결합할 수 있다.

[0146] 실시형태에서, 본 개시내용에 따른 항-백혈병(예컨대, 항-AML) TCR은 서열번호 191 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 상보성 결정 영역 3(CDR3)을 포함하는 TCR베타(β) 사슬을 포함한다.

[0147] 실시형태에서, TCR은 하기 TAP: SLLSGLLRA, ALPVALPSL, ALDPLLLRI IASPIALL 및/또는 SLDLLPLSI 중 하나 이상에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP SLLSGLLRA에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP ALPVALPSL에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP ALDPLLLRI에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP IASPIALL에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP SLDLLPLSI에 특이적이고, 서열번호 191 내지 199에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다.

[0148] 다른 실시형태에서, TCR은 하기 TAP: LTDRIYLTL, VLFGGKVSGA, LGISLTLKY, FNVALNARY 및/또는 TLNQGINVYI 중 하나 이상에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP LTDRIYLTL에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP VLFGGKVSGA에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP LGISLTLKY에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP FNVALNARY에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포

함한다. 실시형태에서, TCR은 TAP TLNQGINVYI에 특이적이고, 서열번호 200 내지 209에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다.

[0149] 다른 실시형태에서, TCR은 하기 TAP: LRSQILSY, KILDVNLRI, HSLISIVYL, KLQDKEIGL 및/또는 AQDIILQAV 중 하나 이상에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP LRSQILSY에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP KILDVNLRI에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP HSLISIVYL에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은TAP KLQDKEIGL에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다. 실시형태에서, TCR은 TAP AQDIILQAV에 특이적이고, 서열번호 210 내지 219에 제시된 아미노산 서열 중 하나를 포함하는 CDR3을 포함하는 TCRβ 사슬을 포함한다.

[0150] 실시형태에서, 본 개시내용에 따른 TCR은 HLA-A*02:01, HLA-A*29:02, HLA-B*15:01, HLA-B27:05, HLA-C*01:02, 및/또는 HLA-C*03:04 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA-A*02:01 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA-A*29:02 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA -B*15:01 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA-B27:05 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA-C*01:02 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다. 실시형태에서, 본 개시내용에 따른 TCR은 HLA-C*03:04 분자에 결합된 위에서 언급된 TAP 중 하나 이상을 인식한다.

[0151] 본 명세서에서 사용되는 바와 같은 용어 TCR은, MHC 수용체에 결합된 항원 펩타이드에 특이적으로 결합 가능한 가변 결합 도메인, 불변 도메인, 막관통 영역 및 짧은 세포질 꼬리를 갖는 면역글로불린 수퍼패밀리 구성원을 지칭한다(예컨대, 문헌[Janeway et al, Immunobiology: The Immune System in Health and Disease, 3rd Ed., Current Biology Publications, p. 4:33, 1997] 참조). TCR은 세포의 표면 상에서 발견될 수 있고, 일반적으로 α 및 β 사슬을 갖는 헤테로다이머(각각 TCRα 및 TCRβ로고 공지됨)로 구성된다. 면역글로불린과 마찬가지로, TCR 사슬(예컨대, α-사슬, β-사슬)의 세포외 부분은 2개의 면역글로불린 영역, 가변 영역(예컨대, TCR 가변 α 영역 또는 Vα 및 TCR 가변 β 영역 또는 Vβ; 전형적으로 N-말단에서 번호매김하는 Rabat에 기반한 아미노산 1 내지 116), 및 세포막에 인접한 1개의 불변 영역(예컨대, TCR 불변 도메인 α 또는 Cα 및 전형적으로 Rabat에 기반한 아미노산 117 내지 259, TCR 불변 도메인 β 또는 Cβ, 전형적으로 Rabat에 기반한 아미노산 117 내지 295)을 함유한다. 또한, 면역글로불린과 마찬가지로, 가변 도메인은 프레임워크 영역(FR)에 의해 분리된 상보성 결정 영역(각 사슬에서 CDR 3)을 함유한다. 소정의 실시형태에서, TCR은 T 세포(또는 T 림프구)의 표면 상에서 발견되고 CD3 복합체와 회합된다.

[0152] TCR 및 특히 본 개시내용의 TCR을 인코딩하는 핵산은, 예를 들어, T 림프구(예컨대, $CD8^+$ T 림프구) 또는 MHC 클래스 I/TAP 복합체를 특이적으로 인식하는 새로운 T 림프구 클론을 생성하는 다른 유형의 림프구를 유전자 형질전환/변형시키는 데 적용될 수 있다. 특정 실시형태에서, 환자로부터 얻어진 T 림프구(예컨대, $CD8^+$ T 림프구)는 TAP를 인식하는 하나 이상의 TCR을 발현하도록 형질전환되고, 형질전환된 세포는 환자에게 투여된다(자가 세포 수혈). 특정 실시형태에서, 공여체로부터 얻어진 T 림프구(예컨대, $CD8^+$ T 림프구)는 TAP를 인식하는 하나 이상의 TCR을 발현하도록 형질전환되고, 형질전환된 세포는 수용체에게 투여된다(동종 세포 수혈). 다른 실시형태에서, 본 개시내용은 TAP-특이적 TCR을 인코딩하는 벡터 또는 플라스미드에 의해 형질전환된/형질감염된 T 림프구, 예컨대, $CD8^+$ T 림프구를 제공한다. 추가의 실시형태에서, 본 개시내용은 TAP-특이적 TCR과 형질전환된 자가 또는 동종 세포로 환자를 치료하는 방법을 제공한다. 소정의 실시형태에서, TCR은, 예컨대, CRISPR, TALEN, 아연 집게(zinc finger), 또는 기타 표적화된 파쇄 시스템을 사용하여, 내인성 좌위, 예컨대, 내인성 TRAC 및/또는 TRBC 좌위를 대체함으로써 1차 T 세포(예컨대, 세포독성 T 세포)에서 발현된다.

[0153] 다른 실시형태에서, 본 개시내용은 위에서 언급된 TCR을 인코딩하는 핵산을 제공한다. 추가의 실시형태에서, 핵산은 벡터, 예컨대, 위에서 기재된 벡터에 존재한다.

[0154] 더욱 추가의 실시형태에서, 암(AML과 같은 백혈병)을 치료하기 위한 자가 또는 동종 세포의 제조에서의 종양 항원-특이적 TCR의 용도가 제공된다.

[0155] 일부 실시형태에서, 본 개시내용의 조성물(예컨대, 약제학적 조성물)로 치료된 환자는 동종 줄기 세포 이식(allogenic stem cell transplant: ASCL), 동종 림프구 주입 또는 자가 림프구 주입으로 치료 전에 또는 치료 후에 치료된다. 본 개시내용의 조성물은 TAP에 대해서 생체외에서 활성화된 동종 T 림프구(예컨대, $CD8^+$ T 림프구); TAP가 로딩된 동종 또는 자가 APC 백신; TAP 백신 및 동종 또는 자가 T 림프구(예컨대, $CD8^+$ T 림프구) 또는 종양 항원-특이적 TCR로 형질전환된 림프구를 포함한다. 본 개시내용에 따른 TAP를 인식 가능한 T 림프구 클론을 제공하는 방법은 대상체(예컨대, 이식편 수용체), 예를 들어, ASCT 및/또는 공여체 림프구 주입(donor lymphocyte infusion: DLI) 수용체에서 TAP를 발현하는 종양 세포에 대해서 발생될 수 있고, 이에 특이적으로 표적화될 수 있다. 따라서, 본 개시내용은 TAP/MHC 클래스 I 분자 복합체를 특이적으로 인식 또는 결합할 수 있는 T 세포 수용체를 인코딩하고 발현하는 $CD8^+$ T 림프구를 제공한다. 상기 T 림프구(예컨대, $CD8^+$ T 림프구)는 재조합(조작된) 또는 천연적으로 선택된 T 림프구일 수 있다. 따라서, 본 명세서는, 본 개시내용의 $CD8^+$ T 림프구를 제조하는 적어도 2가지 방법을 제공하되, 해당 방법은 T 세포 활성화 및 확장을 촉발시키는 데 도움이 되는 조건하에서 미분화된 림프구를 (전형적으로, APC와 같은 세포의 표면에서 발현된) TAP/MHC 클래스 I 분자 복합체와 접촉시키는 단계를 포함하고, 이는 시험관내에서 또는 생체내에서(즉, APC에 TAP가 로딩된 APC 백신을 투여받은 환자에서, 또는 TAP 백신으로 치료된 환자에서) 수행될 수 있다. MHC 클래스 I 분자에 결합된 TAP의 조합물 또는 풀을 이용하여, 복수의 TAP를 인식할 수 있는 $CD8^+$ T 림프구 집단을 생성할 수 있다. 대안적으로, 종양 항원-특이적 또는 표적화된 T 림프구는 MHC 클래스 I 분자/TAP 복합체(즉, 조작된 또는 재조합 $CD8^+$ T 림프구)에 특이적으로 결합하는 TCR(더욱 구체적으로, 알파 및 베타 사슬)을 인코딩하는 하나 이상의 핵산(유전자)을 클로닝함으로써 시험관내에서 또는 생체외에서 제조/생성될 수 있다. 본 개시내용의 TAP-특이적 TCR을 인코딩하는 핵산은, 당업계에 공지된 방법을 사용하여 (예컨대, TAP가 로딩된 APC를 이용하여) 생체외에서 TAP에 대해 활성화된 T 림프구로부터; 또는 펩타이드/MHC 분자 복합체에 대해 면역 반응을 나타내는 개체로부터 얻어질 수 있다. 본 개시내용의 TAP-특이적 TCR은 이식편 수용체 또는 이식편 공여체로부터 얻어진 숙주 세포 및/또는 숙주 림프구에서 재조합적으로 발현될 수 있고, 선택적으로, 시험관내에서 분화되어 세포독성 T 림프구(cytotoxic T lymphocyte: CTL)를 제공할 수 있다. TCR 알파 및 베타 사슬을 인코딩하는 핵산(들)(트랜스진(transgene)(들))은 임의의 적합한 방법, 예컨대, 형질감염(예컨대, 전기천공) 또는 형질도입(예컨대, 바이러스 벡터 사용), 예컨대, 인산칼슘-DNA 공침, DEAE-덱스트란-매개 형질감염, 폴리브렌-매개 형질감염, 전기천공, 미세주입, 리포솜 융합, 리포펙션, 원형질체 융합, 레트로바이러스 감염, 및 유전자총(biolistics)을 사용하여, (예컨대, 치료하고자 하는 대상체 또는 또 다른 개체로부터의) T 세포 내로 도입될 수 있다. TAP에 특이적인 TCR을 발현하는 조작된 $CD8^+$ T 림프구는 널리 공지된 배양 방법을 이용하여 시험관내에서 확장될 수 있다.

[0156] 본 개시내용은 본 명세서에 기재된 바와 같은 TCR을 발현하는 면역 효과기 세포를 제조하기 위한 방법을 제공한다. 일 실시형태에서, 방법은 면역 효과기 세포, 예컨대, 대상체, 예컨대, 백혈병(예컨대, AML)을 지닌 대상체로부터 단리된 면역 효과기 세포를 형질감염 또는 형질도입하는 것을 포함하므로, 면역 효과기 세포는 본 명세서에 기재된 바와 같은 하나 이상의 TCR을 발현하게 한다. 소정의 실시형태에서, 면역 효과기 세포는 개체로부터 단리되고 시험관내에서 추가의 조작 없이 유전자 변형된다. 이어서, 이러한 세포는 개체에 직접 투여될 수 있다. 추가의 실시형태에서, 면역 효과기 세포는 먼저 활성화되고 자극되어, TCR을 발현하도록 유전자 변형되기 전에 시험관내에서 증식된다. 이와 관련하여, 면역 효과기 세포는 유전자 변형되기(즉, 본 명세서에 기재된 바와 같이 TCR를 발현하도록 형질도입 또는 형질감염되기) 전에 또는 후에 배양될 수 있다.

[0157] 본 명세서에 기재된 면역 효과기 세포의 시험관내 조작 또는 유전자 변형 전에, 세포의 공급원은 대상체로부터 얻어질 수 있다. 특히, 본 명세서에 기재된 바와 같은 TCR과 사용하기 위한 면역 효과기 세포는 T 세포를 포함한다. T 세포는 말초 혈액 단핵세포(PBMC), 골수, 림프절 조직, 제대혈, 흉선 조직, 감염 부위로부터의 조직, 복수, 흉수, 비장 조직 및 종양을 비롯한 많은 공급원으로부터 얻어질 수 있다. 소정의 실시형태에서, T 세포는 FICOLL™ 분리와 같은 당업자에게 공지된 임의의 수의 기술을 사용하여 대상체로부터 수집된 혈액 단위로부터 얻어질 수 있다. 일 실시형태에서, 개체의 순환 혈액으로부터의 세포는 성분채집에 의해 얻어진다. 성분채집 제품은 전형적으로 T 세포, 단핵구, 과립구, B 세포, 기타 유핵 백혈구 세포, 적혈구 세포 및 혈소판을 비롯한 림프구를 함유한다. 일 실시형태에서, 성분채집에 의해 수집된 세포는 혈장 분획을 제거하고, 후속 처리를 위해 적절한 완충액 또는 배지에 세포를 배치하기 위해 세척될 수 있다. 본 발명의 일 실시형태에서, 세포는 PBS로 세척된다. 대안적인 실시형태에서, 세척된 용액은 칼슘이 부족하고, 마그네슘이 부족할 수 있거나 또는 2가 양이온이 전부는 아니더라도 많이 부족할 수 있다. 당업자에 의해 이해되는 바와 같이, 세척 단계는 당업자에게

공지된 방법에 의해, 예컨대, 반자동 관류 원심분리기를 사용하는 것에 의해 달성될 수 있다. 세척 후, 세포는 다양한 생체적합성 완충액 또는 완충액이 있거나 없는 기타 염수 용액에 재현탁될 수 있다. 소정의 실시형태에서, 성분채집 샘플의 바람직하지 않은 성분은 세포가 직접 재현탁된 배양 배지에서 제거될 수 있다. 소정의 실시형태에서, T 세포는, 적혈구를 용해시키고 단핵구를 고갈시킴으로써, 예를 들어, PERCOLL™ 구배를 통한 원심분리에 의해 말초 혈액 단핵 세포(PBMC)로부터 단리된다. CD28+, CD4+, CD8+, CD45RA+ 및 CD45RO+ T 세포와 같은 T 세포의 특정 하위집단은 양성 또는 음성 선택 기술에 의해 더욱 단리될 수 있다. 예를 들어, 음성 선택에 의한 T 세포 집단의 농축은 음성 선택 세포에 고유한 표면 마커에 지향된 항체의 조합으로 달성될 수 있다. 본 명세서에서 사용하기 위한 한 가지 방법은 음성으로 선택된 세포에 존재하는 세포 표면 마커에 지향된 단클론성 항체의 칵테일을 사용하는 음성 자기 면역부착 또는 유세포분석을 통한 세포 분류 및/또는 선택이다. 예를 들어, 음성 선택에 의해 CD8+ 세포를 강화하기 위해, 단클론성 항체 칵테일은 전형적으로 CD14, CD20, CD11b, CD16, HLA-DR 및 CD4에 대한 항체를 포함한다. 유세포분석 및 세포 분류는 또한 본 개시내용에서 사용하기 위한 관심 세포 집단을 단리시키는 데 사용될 수 있다. PBMC는 본 명세서에 기재된 바와 같은 방법을 사용하여 TCR로 유전자 변형시키기 위하여 직접 사용될 수 있다. 소정의 실시형태에서, PBMC의 단리 후, T 림프구는 추가로 단리되고, 소정의 실시형태에서, 세포독성 및 헬퍼 T 림프구는 둘 다 유전자 변형 및/또는 확장 전에 또는 후에 미경험, 기억 및 효과기 T 세포 하위집단으로 분류될 수 있다.

[0158] 본 개시내용은 TAP(즉, 세포의 표면에서 발현된 MHC 클래스 I 분자에 결합된 TAP) 또는 TAP의 조합물에 의해 특이적으로 유도되고, 활성화되고 그리고/또는 증폭된(확장된) $CD8^+$ T 림프구와 같은 단리된 면역 세포를 제공한다. 본 개시내용은 또한 본 개시내용에 따른 TAP 또는 이의 조합물(즉, MHC 클래스 I 분자에 결합된 하나 이상의 TAP) 및 상기 TAP(들)를 인식 가능한 $CD8^+$ T 림프구를 포함하는 조성물을 제공한다. 다른 양상에서, 본 개시내용은 본 명세서에 기재된 바와 같은 하나 이상의 MHC 클래스 I 분자/TAP 복합체(들)를 특이적으로 인식하는 $CD8^+$ T 림프구가 농축된 세포 집단 또는 세포 배양물(예컨대, $CD8^+$ T 림프구 집단)을 제공한다. 이러한 농축된 집단은, 본 명세서에 개시된 바와 같은 TAP 중 하나 이상의 것으로 로딩된(예컨대, 이를 제시하는) MHC 클래스 I 분자를 발현하는 APC와 같은 세포를 이용하여 특이적인 T 림프구의 생체외 확장을 수행함으로써 얻어질 수 있다. 본 명세서에서 사용되는 바와 같은 "농축된(enriched)"은, 집단 중 종양 항원-특이적 $CD8^+$ T 림프구의 비율이 천연 세포 집단에 비하여 유의하게 더 높은 것을 의미하며, 즉, 특이적인 T 림프구의 생체외 확장 단계를 거치지 않았음을 의미한다. 추가의 실시형태에서, 세포 집단 중 TAP-특이적 $CD8^+$ T 림프구의 비율은 적어도 약 0.5%, 예를 들어 적어도 약 1%, 1.5%, 2% 또는 3%이다. 일부 실시형태에서, 세포 집단 중 TAP-특이적 $CD8^+$ T 림프구의 비율은 약 0.5 내지 약 10%, 약 0.5 내지 약 8%, 약 0.5 내지 약 5%, 약 0.5 내지 약 4%, 약 0.5 내지 약 3%, 약 1% 내지 약 5%, 약 1% 내지 약 4%, 약 1% 내지 약 3%, 약 2% 내지 약 5%, 약 2% 내지 약 4%, 약 2% 내지 약 3%, 약 3% 내지 약 5% 또는 약 3% 내지 약 4%이다. 하나 이상의 관심 MHC 클래스 I 분자/펩타이드(TAP) 복합체(들)를 특이적으로 인식하는 $CD8^+$ T 림프구가 농축된 이러한 세포 집단 또는 배양물(예컨대, $CD8^+$ T 림프구 집단)은 이하에서 상세히 설명되는 바와 같은 종양 항원-기반 암 면역요법에서 사용될 수 있다. 일부 실시형태에서, TAP-특이적 $CD8^+$ T 림프구의 집단은, 예를 들어, 친화도-기반 시스템, 예컨대, 본 명세서에서 정의된 TAP(들)가 (공유적으로 또는 공유적이 아니게) 로딩된 MHC 클래스 I 분자의 다량체를 이용하여 추가로 농축된다. 따라서, 본 개시내용은, 예컨대, TAP-특이적 $CD8^+$ T 림프구의 비율이 적어도 약 50%, 60%, 70%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, 99% 또는 100%인 TAP-특이적 $CD8^+$ T 림프구의 정제된 또는 단리된 집단을 제공한다.

[0159] 본 개시내용은 추가로 위에서 언급된 면역 세포($CD8^+$ T 림프구) 또는 TAP-특이적 $CD8^+$ T 림프구의 집단을 포함하는 약제학적 조성물 또는 백신에 관한 것이다. 이러한 약제학적 조성물 또는 백신은 위에서 기재된 바와 같은 1종 이상의 약제학적으로 허용 가능한 부형제 및/또는 애주번트를 포함할 수 있다.

[0160] 본 개시내용은 추가로 의약으로서의, 또는 의약의 제조에서의, 본 개시내용에 따른 임의의 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물, 또는 이들의 임의의 조합물의 용도에 관한 것이다. 실시형태에서, 의약은 암, 예컨대, 암 백신의 치료용이다. 본 개시내용은 암 치료에서, 예컨대, 암 백신으로서 사용하기 위한, 본 개시내용에 따른 임의의 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물(예컨대, 백신 조성물), 또는 이들의 임의의 조합물에 관한 것이다. 본 명세서에서 식별된 TAP 서열은 i) 종양 환자 내로 주사된 종양 항원-특이적 T 세포의 시험관내 프라이밍 및 확장에 사용되

고/되거나, ii) 암 환자에서 항종양 T 세포 반응을 유도 또는 부스팅하는 백신으로서 사용될 합성 펩타이드의 제조에 사용될 수 있다.

[0161] 다른 양상에서, 본 개시내용은, 대상체에서 암을 치료하기 위한 백신으로서의, 본 명세서에 기재된 TAP(서열번호 1 내지 190, 바람직하게는 서열번호 97 내지 154), 또는 이들의 조합물(예컨대, 펩타이드 풀)의 용도를 제공한다. 본 개시내용은 또한 대상체에서 암을 치료하기 위한 백신으로서 사용하기 위한, 본 명세서에 기재된 TAP, 또는 이들의 조합물(예컨대, 펩타이드 풀)을 제공한다. 실시형태에서, 대상체는 TAP-특이적 $CD8^+$ T 림프구의 수용체이다. 따라서, 다른 양상에서, 본 개시내용은 암을 치료하는(예컨대, 종양 세포의 수를 감소시키는, 종양 세포를 사멸시키는) 방법을 제공하되, 상기 방법은 암 치료를 필요로 하는 대상체에게 (예컨대, APC와 같은 세포의 표면에서 발현된) 하나 이상의 MHC 클래스 I 분자/TAP 복합체를 인식하는(즉, 이에 결합하는 TCR을 발현하는) 유효량의 $CD8^+$ T 림프구를 투여하는(주입하는) 단계를 포함한다. 실시형태에서, 방법은 상기 $CD8^+$ T 림프구의 투여/주입 후 상기 대상체에게 유효량의 TAP, 또는 이들의 조합물, 및/또는 TAP(들)가 로딩된 MHC 클래스 I 분자(들)를 발현하는 세포(예컨대, APC, 예컨대, 수지상 세포)를 투여하는 단계를 더 포함한다. 더욱 추가의 실시형태에서, 방법은 이를 필요로 하는 대상체에게 하나 이상의 TAP가 로딩된 치료적 유효량의 수지상 세포를 투여하는 단계를 포함한다. 더욱 추가의 실시형태에서, 방법은 이를 필요로 하는 환자에게 MHC 클래스 I 분자에 의해 제시된 TAP에 결합하는 재조합 TCR을 발현하는 치료적 유효량의 동종 또는 자가 세포를 투여하는 단계를 포함한다.

[0162] 다른 양상에서, 본 개시내용은 대상체에서 암을 치료하기 위한(예컨대, 종양 세포를 감소시키기 위한, 종양 세포를 사멸시키기 위한), TAP 또는 이의 조합물이 로딩된 (이를 제시하는) 하나 이상의 MHC 클래스 I 분자를 인식하는 $CD8^+$ T 림프구의 용도를 제공한다. 다른 양상에서, 본 개시내용은 대상체에서 암을 치료하기 위한(예컨대, 종양 세포를 감소시키기 위한, 종양 세포를 사멸시키기 위한) 의약의 제조/생산을 위한, TAP 또는 이의 조합물이 로딩된 (이를 제시하는) 하나 이상의 MHC 클래스 I 분자를 인식하는 $CD8^+$ T 림프구의 용도를 제공한다. 다른 양상에서, 본 개시내용은 대상체에서의 암 치료에서 사용하기 위한(예컨대, 종양 세포를 감소시키기 위한, 종양 세포를 사멸시키 위한), TAP 또는 이의 조합물이 로딩된 (이를 제시하는) 하나 이상의 MHC 클래스 I 분자(들)를 인식하는 $CD8^+$ T 림프구(세포독성 T 림프구)를 제공한다. 추가의 실시형태에서, 용도는 상기 TAP-특이적 $CD8^+$ T 림프구의 사용 후, 유효량의 TAP(또는 이의 조합물), 및/또는 TAP가 로딩된 (이를 제시하는) 하나 이상의 MHC 클래스 I 분자(들)를 발현하는 세포(예컨대, APC)의 용도를 더 포함한다.

[0163] 본 개시내용은 또한 대상체에서 본 명세서에 개시된 TAP 중 임의의 것 또는 이의 조합물이 로딩된 인간 클래스 I MHC 분자를 발현하는 종양 세포(백혈병 세포, AML 세포)에 대해 면역 반응을 생성하는 방법을 제공하되, 해당 방법은 TAP 또는 TAP의 조합물이 로딩된 클래스 I MHC 분자를 특이적으로 인식하는 세포독성 T 림프구를 투여하는 단계를 포함한다. 본 개시내용은 또한 TAP 또는 이의 조합물이 로딩된 인간 클래스 I MHC 분자를 발현하는 종양 세포에 대해 면역 반응을 생성하기 위한, 본 명세서에 개시된 TAP 중 임의의 것 또는 TAP의 조합물이 로딩된 클래스 I MHC 분자를 특이적으로 인식하는 세포독성 T 림프구의 용도를 제공한다.

[0164] 실시형태에서, 본 명세서에 기재된 방법 또는 용도는 치료/사용 전에 환자에 의해 발현되는 HLA 클래스 I 대립유전자를 결정하는 단계, 및 환자에 의해 발현되는 HLA 클래스 I 대립유전자 중 하나 이상에 결합하는 TAP를 투여 또는 사용하는 단계를 더 포함한다. 예를 들어, 환자가 HLA-A1*01 및 HLA-C05*01을 발현하는 것으로 결정된 경우, (i) (HLA-A1*01에 결합하는) 서열번호 48, 67, 89, 134, 151 및/또는 164, 및/또는 (HLA-C05*01에 결합하는) 서열번호 150의 TAP의 임의의 조합물이 환자에게 투여될 수 있거나 환자에서 사용될 수 있다.

[0165] 실시형태에서, 암은 혈액암, 바람직하게는, 백혈병, 예컨대, 급성 림프구 백혈병(ALL), 급성 골수성 백혈병(AML), 만성 골수성 백혈병(CML), 모발세포 백혈병(HCL) 및 골수이형성 증후군(MDS)이다. 실시형태에서, 백혈병은 AML이다. 본 명세서에 기재된 방법 및 사용에 의해 치료된 AML은, 임의의 유형 또는 아형(예컨대, 저-, 중간- 또는 고-위험 AML), 예를 들어, 유전적 기형(genetic abnormalities)을 가진 AML, 예컨대, 염색체 8과 21 간의 전좌[t(8;21)]를 가진 AML, 염색체 16에서의 전좌 또는 역위[t(16;16) 또는 inv(16)]를 가진 AML, PML-RARA 융합 유전자를 가진 AML, 염색체 9와 11 사이의 전좌[t(9;11)]를 가진 AML, 염색체 6과 9 사이의 전좌[t(6:9)]를 가진 AML, 염색체 3에서의 전좌 또는 역위[t(3;3) 또는 inv(3)]를 가진 AML, 염색체 1과 22 사이의 전좌[t(1:22)]를 가진 AML(거대핵모구성), BCR-ABL1(BCR-ABL) 융합 유전자를 가진 AML, 돌연변이된 NPM1 유전자를 가진 AML, CEBPA 유전자의 이중대립유전자 돌연변이를 가진 AML, 돌연변이된 RUNX1 유전자를 가진 AML, 돌연변이된 ASX1 유전자를 가진 AML, 돌연변이된 IDH1 및/또는 IDH2 유전자를 가진 AML, 돌연변이된 FLT3 유전자를 가

진 AML, 골수이형성-관련 변화를 가진 AML, 및 이전의 화학요법 또는 방사선과 관련된 AML일 수 있다.

[0166] 실시형태에서, 본 개시내용에 따른 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물, 또는 이들의 임의의 조합물은, 암을 치료하는 하나 이상의 추가의 활성제 또는 요법, 예컨대, 화학요법(예컨대, 빈카 알카로이드, 미세소관 형성을 파괴하는 제제(예컨대, 콜히친 및 이의 유도체), 항혈관신생제, 치료용 항체, EGFR 표적화제, 티로신 키나제 표적화제(예컨대, 티로신 키나제 저해제), 전이금속 복합체, 프로테아좀 저해제, 항대사산물(예컨대, 뉴클레오사이드 유사체), 알킬화제, 백금-기반 제제, 안트라사이클린 항생제, 토포아이소머라제 저해제, 매크로라이드, 레티노이드(예컨대, 올-트랜스(all-trans) 레티노산 또는 이의 유도체), 겔다나마이신 또는 이의 유도체(예컨대, 17-AAG), 수술, 방사선요법, 면역관문저해제(면역요법제(예컨대, PD-1/PD-L1 저해제, 예컨대, 항-PD-1/PD-L1 항체, CTLA-4 저해제, 예컨대, 항-CTLA-4 항체, B7-1/B7-2 저해제, 예컨대, 항-B7-1/B7-2 항체, TIM3 저해제, 예컨대, 항-TIM3 항체, BTLA 저해제, 예컨대, 항-BTLA 항체, CD47 저해제, 예컨대, 항-CD47 항체, GITR 저해제, 예컨대, 항-GITR 항체), 종양 항원에 대한 항체, 세포-기반 요법(예컨대, CAR T 세포, CAR NK 세포), 사이토카인, 예컨대, IL-2, IL-7, IL-21 및 IL-15와 조합하여 사용될 수 있다. 실시형태에서, 본 개시내용에 따른 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물은 면역관문저해제와 조합하여 투여/사용된다. 실시형태에서, 본 개시내용에 따른 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물은 AML의 치료에 사용되는 하나 이상의 화학치료 약물과 조합하여, 또는 기타 AML 요법, 예를 들어, 줄기세포/골수 이식과 조합하여 투여/사용된다.

[0167] 추가의 요법은 본 개시내용에 따른 TAP, 핵산, 발현 벡터, T 세포 수용체, 세포(예컨대, T 림프구, APC) 및/또는 조성물 전에, 이와 동시에 또는 이후에 투여될 수 있다.

[0168] **발명을 수행하기 위한 방식(들)**

[0169] 본 발명은 이하의 비제한적인 실시예에 의해 더욱 상세히 예시된다.

[0170] **실시예 1: 재료 및 방법**

[0171] *AML 시편*

[0172] 진단용 AML 샘플(DMSO-동결된 백혈병 모세포의 냉동바이알)은 Banque de cellules leucémiques du Québec 프로그램(BCLQ, bclq.org)으로부터 얻었다. 샘플의 기술적 및 임상적 특징은 **표 1**에 제공된다. 각 AML 샘플(14H124 제외, 하기 부문 참조)의 1억개의 세포를 해동(37℃ 수욕에서 1분)시키고 48㎖의 4℃ PBS에 재현탁시켰다. 2백만개 세포(1㎖)를 펠릿화하고 RNA-시퀀싱을 위하여 1㎖ Trizol에 재현탁시키는 한편, 나머지 9천8백만개를 펠릿화하고, 질량분광분석을 위하여 액체 질소에서 스냅-동결시켰다.

[0173] [표 1]

본 연구에서 TSA 를 식별하는 데 사용되는 19 가지 AML 시편의 생물학적 및 임상적 특징.

[0174]

ID BCLQ	HLA 타이핑 (옵티타입)	WHO 2008	FAB	조직	모세포%	RNA-seq 리드 계수치

05H143	A*02:01, A*34:02, B*08:01, B*57:01, C*07:01, C*06:02	급성 단핵구성 및 단구성 백혈병	M5A	말초 혈액	85	2,14E+08
05H149	A*24:02, A*68:02, B*53:01, B*18:01, C*04:01, C*05:01	급성 골수성 백혈병, NOS	M1	말초 혈액	80	1,19E+08
07H060	A*01:01, A*29:02, B*08:01, B*56:01, C*07:01, C*01:02	성숙하지 않은 AML	M1	말초 혈액	90	2,04E+08
07H063	A*30:01, A*01:01, B*13:02, B*07:02, C*06:02, C*07:02	요법-관련 골수성 신생물	M5A	골수	96	1,92E+08
07H122	A*02:01, A*02:01, B*57:01, B*57:01, C*06:02, C*06:02	골수이형성-관련 변화를 가진 AML	NC	말초 혈액	96	1,84E+08
07H141	A*02:01, A*26:01, B*38:01, B*44:02, C*05:01, C*12:03	성숙하지 않은 AML	M1	말초 혈액	75	1,77E+08
08H039	A*02:01, A*01:01, B*40:01, B*14:02, C*03:04, C*08:02	최소 분화를 가진 AML	M0	말초 혈액	96	1,82E+08
08H053	A*01:01, A*11:01, B*14:02, B*15:01, C*03:03, C*03:03	성숙하지 않은 AML	M1	말초 혈액	95	1,74E+08
10H005	A*11:01, A*11:01, B*13:01, B*13:01, C*07:02, C*03:04	급성 골수단구성 백혈병	M4	골수	71	1,84E+08
11H008	A*29:02, A*03:01, B*44:03, B*57:01, C*06:02, C*16:01	골수이형성-관련 변화를 가진 AML	NC	말초 혈액	75	1,76E+08
11H035	A*03:01, A*24:02, B*07:02, B*07:02, C*07:02, C*07:02	성숙하지 않은 AML	M1	말초 혈액	93	1,80E+08
12H172	A*24:02, A*02:01, B*15:01, B*15:01, C*03:03, C*03:03	요법-관련 골수성 신생물	M1	말초 혈액	85	1,91E+08
14H124	A*02:01, A*26:01, B*44:05, B*27:05, C*02:02, C*02:02	성숙 AML	M2	말초 혈액	86	1,23E+08
15H013	A*30:02, A*02:01, B*39:05, B*57:03, C*18:01, C*07:02	성숙하지 않은 AML	M1	말초 혈액	90	1,61E+08
15H023	A*02:01, A*03:01, B*51:01, B*44:03, C*01:02, C*02:02	성숙하지 않은 AML	M1	말초 혈액	90	1,79E+08
15H063	A*02:01, A*01:01, B*08:01, B*40:01, C*07:01, C*03:04	골수이형성-관련 변화를 가진 AML	NC	말초 혈액	85	1,65E+08
15H080	A*02:01, A*26:01, B*44:05, B*27:05, C*02:02, C*02:02	급성 단핵구성 및 단구성 백혈병	M5A	말초 혈액	70	1,95E+08
16H123	A*24:02, A*03:01, B*15:01, B*07:02, C*07:02, C*03:03	성숙하지 않은 AML	M1	말초 혈액	88	1,66E+08
16H145	A*02:01, A*02:01, B*27:02, B*55:01, C*03:03, C*02:02	최소 분화를 가진 AML	M0	말초 혈액	88	1,85E+08

[0175]

ID BCLQ	고유한 k-mer (33) 계수치	mTEC 고갈에서 사용되는 발생 역치	필터링 발생 후의 k-mer 계수치	mTEC 필터링 후의 k-mer 계수치	mTEC+MPC 고갈에서 사용되는 발생 역치	필터링 발생 후의 k-mer 계수치
05H143	1,44E+09	10	1,26E+08	3,01E+07	3	3,38E+08
05H149	8,89E+08	12	7,64E+07	2,33E+07	5	1,78E+08
07H060	1,24E+09	8	1,11E+08	2,02E+07	3	2,76E+08
07H063	9,55E+08	6	9,26E+07	1,62E+07	3	1,80E+08
07H122	9,67E+08	7	8,35E+07	1,85E+07	3	2,03E+08
07H141	1,08E+09	10	8,81E+07	1,55E+07	3	2,55E+08
08H039	1,10E+09	8	9,72E+07	1,85E+07	3	2,47E+08
08H053	1,14E+09	8	9,08E+07	1,56E+07	3	2,40E+08
10H005	1,01E+09	6	8,63E+07	1,56E+07	3	1,81E+08
11H008	8,30E+08	5	9,73E+07	1,79E+07	3	1,68E+08
11H035	1,07E+09	7	1,06E+08	1,86E+07	3	2,39E+08
12H172	1,12E+09	6	1,07E+08	2,17E+07	3	2,18E+08
14H124	7,77E+08	4	1,03E+08	1,76E+07	3	1,35E+08
15H013	8,65E+08	5	1,03E+08	1,86E+07	3	1,73E+08
15H023	1,21E+09	8	1,12E+08	2,28E+07	3	2,75E+08
15H063	1,02E+09	8	9,79E+07	2,16E+07	3	2,29E+08
15H080	1,13E+09	7	1,11E+08	2,23E+07	3	2,44E+08
16H123	1,06E+09	8	1,01E+08	1,77E+07	3	2,50E+08
16H145	1,19E+09	8	1,08E+08	2,12E+07	3	2,65E+08

[0176]

ID BCLQ	mTEC + MPC 필터링 후의 k-mer 계수치	특유의 MAP 계수치			
		mTEC 고갈	ERE	mTEC+MPC 고갈	k-mer 발현량 차이
05H143	2,53E+07	2894	2358	2897	2877
05H149	2,47E+07	542	519	545	539
07H060	1,83E+07	986	957	983	973
07H063	8,14E+06	1784	1777	1797	1715
07H122	2,53E+07	2940	2702	2911	2878
07H141	1,37E+07	1662	1554	1647	1544
08H039	2,27E+07	2594	2420	2689	2670
08H053	1,81E+07	317	270	328	292
10H005	1,06E+07	220	136	280	159
11H008	9,91E+06	1163	1054	1175	1072
11H035	1,53E+07	1928	1899	1986	1933
12H172	1,59E+07	2865	2916	2861	2851
14H124	3,63E+06	2160	2093	2249	2116
15H013	8,86E+06	1761	1686	1756	1748
15H023	1,68E+07	1794	1733	1799	1732
15H063	1,25E+07	1277	1128	1285	1190
15H080	1,23E+07	797	722	796	760
16H123	1,80E+07	1424	1383	1467	1391
16H145	1,71E+07	755	704	765	742

NC: FAB 기준에 의해 분류 가능하지 않음. HLA 는 각 샘플의 RNA-Seq 데이터에 기초한 옵티타입에 의해 결정되었다. 임상 데이터는 Banque de cellules leucémiques du Québec program (BCLQ, bclq.org)에 의해 제공되었다.

[0177]

[0178] *데이터의 기타 출처*

[0179] 인간 mTEC 샘플은 본 발명자들의 팀의 이전의 연구(#GSE127825 및 # GSE127826)(Larouche *et al.*, 2020; Laumont *et al.*, 2018)를 위하여 준비되고 시퀀싱되었거나, 또는 다른 것(E-MTAB-7383)(Fergusson *et al.*, 2018)에 의해 간행되었다. 본 발명자들의 군에 의한 TSA 발견을 위하여 이전에 사용된 6개의 mTEC 샘플만이 k-mer 고갈 접근법(Laumont *et al.*, 2018)에서 사용되었다. 주된 정상 대조군으로서 사용된 11개의 MPC 샘플은 IRIC 게놈 플랫폼에 의해 시퀀싱되었고 Leucegene 그룹(#GSE98310, #GSE51984)에 의해 이전에 간행되었다. 본 연구에서 사용된 모든 다른 정상 샘플은 dbGap(www.ncbi.nlm.nih.gov/gap/), Arrayexpress(www.ebi.ac.uk/arrayexpress/) 또는 GEO(www.ncbi.nlm.nih.gov/geo/)로부터 다운로드되었다. 437개 RNA-시퀀싱된 AML 샘플의 Leucegene 전체 코호트는 발견된 TSA^{hi}의 임상적 유의성을 연구하는 데 사용되었다. RNA-시퀀싱 데이터는 이전에 공개되었고, 개별적으로 입수 가능하다(#GSE49642, #GSE52656, #GSE62190, #GSE66917, #GSE67039)(Lavallee et al., 2015; Macrae et al., 2013; Pabst et al., 2016). 분류된 LSC 및 모세포의 RNA-Seq 데이터는 어딘가에 공개되어 있고, #GSE74246(Corces et al., 2016)으로부터 얻었다. 재발 전 및 재발 후의 AML 모세포의 RNA-seq 데이터(일치된 샘플)는 어딘가에 공개되어 있고(Toffalori et al., 2019), 이러한 샘플의 HLA 유형은 Luca Vago 박사에 의해 친절하게 제공되었다. 외부 출처로부터 얻어진 모든 데이터는 STAR v2.5.1b를 사용하여 GRCh38 게놈 상에 정렬되었다.

[0180] *NSG 마우스에서의 14H124 AML 세포의 확장*

[0181] 이 환자에서 2천만개 세포만 이용 가능하였기 때문에, 환자 14H124의 모세포를 해동하고, PBS로 세척하고, 준치사적인 방사선 조사(2.5Gy 137 Cs-감마 공급원) 후 24시간에 10마리의 NOD-scid IL-2Rɣnull $^{(NSG)}$ 마우스에 정맥내 주사하였다(2×10^6/마우스). 인간 AML 세포 생착은 유세포분석에 의해 122일째에 말초 혈액에서 평가되었다. 간단히 말해서, 100㎕의 혈액을 꼬리 정맥 출혈에 의해 수집하고, RBC 용해 완충제(eBioscience)를 사용하여 적혈구를 고갈시키고, 염색 완충제(PBS+3%FBS)로 세척하고, 항-인간 CD45-PacificBlue(HI30, Biolegend) 및 항-마우스 CD45-PECy5(30-F11, BD)로 4℃에서 20분 동안 염색하고, PBS로 세척하였다. FACS Canto II 유세포분석기(Becton Dickinson) 상에서 데이터를 수집하고, Flowjo® 소프트웨어 7.0(Tree Star Inc., 오리건주 애슐랜드 소재)으로 분석하였다.

[0182] 1%보다 높은 인간 세포 키메라 현상이 8/10 마우스에서 발견되었다. 마우스를 질환의 징후(빈혈, 체중 감소>20% 또는 명백한 종양)에서 이식 후 188 내지 264일 이내에 희생시키고, 골수, 비장 및 고형 종양(견갑간, 목 및 엉덩이 부위에서 또는 신장, 간 및 림프절에서 발견됨)을 수거하였다. 질량 분광분석에 의한 향후의 처리를 위하여 종양을 액체 질소에서 급속 동결시켰다. 비장을 분쇄하고 대퇴골 및 경골(골수의 수집)을 4℃ PBS로 씻어냄

으로써 AML 세포를 수집하였다. 세포에서 적혈구를 고갈시키고, 여과(100㎛)시켜 파편을 제거하고, 계수하고, 순도를 평가하기 위하여 위에서 상세히 기재된 바와 같이 유세포분석을 위하여 처리하거나(5×10^5개 세포), 또는 향후의 RNA 시퀀싱(남아 있는 모든 세포)을 위하여 용해시키고 Trizol®(Invitrogen)에서 동결보존하였다. 1㎤ 미만의 크기를 갖고 순도가 99% 초과인 6백만개의 골수-유래 모세포(**도 14A 내지 B**)가 RNA 시퀀싱을 위하여 이용 가능한 하나의 종양을 선택하여, 질량 분광분석에 의한 MAP 식별을 위하여 처리하였다. 말초 혈액에서 122일째에 인간 키메라 현상(이식 실패)이 없는 2마리의 마우스는 어떠한 질환의 징후도 나타내지 않았고, 실험 말기(264일)에 희석시켰다. 모든 희생은 CO_2 질식에 이은 경추 탈구에 의해 인도적으로 수행되었다. 마우스를 주 3회 질환 징후에 대해서 평가하고 실험 동안 매일 모니터링하였다.

[0183] *RNA 추출, 라이브러리 제조 및 시퀀싱*

[0184] RNA 추출은 RNeasy® Mini 추출 칼럼(Qiagen) 상에 Trizol®/클로로포름 추출 및 정제를 사용하여 수행되었다. 400ng의 총 RNA를 라이브러리 제조에 사용하였다. 총 RNA의 품질을 BioAnalyzer Nano(Agilent)로 평가하였고, 모든 샘플은 8 초과의 RIN을 지녔다. 라이브러리 제조는 KAPA mRNAseq Hyperprep 키트(KAPA, 카탈로그 번호 KK8581)로 수행되었다. 결찰은 Illumina Truseq 지수의 51nM 최종 농도로 이루어졌고, cDNA 라이브러리를 증폭시키기 위하여 12회 PCR 사이클이 필요로 되었다. 샘플 14H124는 4M 세포, 1ug의 총 RNA를 사용하여 별도로 수행하였다. 라이브러리 제조는, 증폭이 12회 대신에 10회 PCR 사이클로 행해진 것을 제외하고 앞서의 샘플과 마찬가지로 수행되었다. 라이브러리는 QuBit 및 BioAnalyzer DNA1000으로 정량화되었다. 모든 라이브러리를 10nM로 희석시키고고, KAPA 라이브러리 정량화 카투(KAPA; 카탈로그 번호 KK4973)를 사용하여 qPCR에 의해 정규화시켰다. 라이브러리는 등몰 농도로 풀링시켰다. 2.8pM의 풀링된 라이브러리를 사용하여 Nextseq High Output Kit 150 사이클(2×80bp)을 이용해서 Illumina Nextseq500으로 시퀀싱을 수행하였다. 샘플당 대략 120 내지 200M 페어드-엔드 PF 리드가 생성되었다. 라이브러리 제조 및 시퀀싱은 IRIC(Institute for Research in Immunology and Cancer's Genomics Platform)에서 수행되었다.

[0185] *샷건 질량 분광분석 식별을 위한 데이터베이스 생성*

[0186] *1) 개인화된 정규 프로테옴의 생성.* 이것은 이전에 상세하게 설명된 바와 같이(Laumont et al., 2018) 수행되었다. 간단히 말해서, RNA-Seq 리드는, Trimmomatic v0.35를 사용하여 트리밍되고, 디폴트 값이 각각 10, 200,000, 200,000 및 "5 -1 5 5"로 대체된 --alignSJoverhangMin, --alignMatesGapMax, --alignIntronMax-- 및 --alignSJstitchMismatchNmax 파라미터를 제외하고 디폴트 파라미터로 실행되는 STAR v2.5.1b(Dobin et al., 2013)를 사용하여 GRCh38.88과 정렬되어, bam 파일을 생성하였다. 최소 대체 계수치 설정이 5인 단일-염기 돌연변이는 freeBayes 1.0.2-16-gd466dde(arXiv:1207.3907)를 사용하여 식별되었다. 전사체 발현은 디폴트 파라미터를 사용하여 kallisto v0.43.0 (Bray et al., 2016)로 백만당 전사체(tpm)로 정량화되었다. 마지막으로, pyGeno를 사용하여, 기준 엑솜에 고품질 샘플-특이적 단일-염기 돌연변이(freeBayes 품질>20)를 삽입하고 발현된 전사체(tpm > 0)에 의해 생성된 알려진 단백질의 샘플-특이적 서열을 엑스포트하여 개인화된 정규 프로테옴의 fasta 파일을 생성하였다.

[0187] 2) *mTEC k-mer 고갈에 의한 AML-특이적 프로테옴의 생성에 의한 AML-특이적 프로테옴(**도 2A**).* 이것은 앞서 상세히 설명된 바와 같이(Laumont et al., 2018) 수행되었다. 간단히 말해서, 각 샘플의 R1 및 R2 fastq 파일은 위에서 보고된 바와 같이 트리밍되고, R1 리드는 FASTX-Toolkit v0.0.14의 fastx_reverse_complement 함수(function)를 사용하여 역보완(reverse complement)되었다. K-MER 데이터베이스(24 또는 33-길이)는 jellyfish v2.2.3(Marcais and Kingsford, 2011)으로 생성하였다. 단일 데이터베이스는 각 AML 샘플에 대해서 생성된 반면 6 mTEC 샘플은 이들의 fastq 파일을 결부시킴으로써 고유한 데이터베이스에서 조합되었다. k-mer 조립의 지속기간(하기 참조)이 3천만개의 k-mer 초과로 기하급수적으로 증가하기 때문에, 각 AML 33-뉴클레오타이드-길이 k-mer 데이터베이스는 조립 단계에 대해서 최대 3천만개의 k-mer에 도달하기 위하여 발생에 대한 샘플-특이적 역치(주어진 k-mer가 데이터베이스에 존재하는 횟수)에 기반하여 필터링되었다(**표 1**). 이 필터링 후, mTEC k-mer 데이터베이스에 적어도 한번 존재하는 k-mer는 각 샘플 데이터베이스로부터 제거되고, 나머지 k-mer는 사내 개발 소프트웨어인 NEKTAR를 사용하여 콘티그로 조립되었다. 간단히 말해서, 제출된 33-뉴클레오타이드-길이 k-mer 중 하나는 동일한 가닥의 32개 뉴클레오티드가 겹치는 연속적인 k-mer로 양쪽 끝에서 확장된 시드로서 무작위로 선택된다(k-mer의 가닥 세트가 사용됨에 따라서, -r 옵션은 비활성화됨). 조립 과정은 k-mer가 조립될 수 없거나 하나 초과의 k-mer가 적합화되는 경우 중지된다(선형 조립의 경우 -1 옵션). 그렇다면, 새로운 시드가 선택되고, 제출된 목록으로부터 모든 k-mer가 한 번 사용될 때까지 조립 과정이 재개된다. 마지막으로, 콘티그

는 사내 파이썬 스크립트를 사용하여 3-프레임 번역되었고, 아미노산 서열은 내부 정지 코돈에서 분할되고, 얻어진 하위서열은 각 샘플 각각의 개인화된 정규 프로테옴과 결부되었다.

[0188] *3) ERE-특이적 프로테옴의 생성(**도 2B**). 각 샘플에 대해서, RNA-Seq* 리드는 디폴트 파라미터와 함께 STAR(Dobin et al., 2013)을 사용하여 인간 참조 게놈(GRCh38.88)에 정렬되었다. BEDtools(PMID 20110278)의 교차 함수(intersect function)를 이용해서, 리드는 ERE 서열 또는 정규 유전자 중 하나에 완전히 매핑되는 2개의 리드 데이터 세트로 분리되었다. ERE 리드 데이터세트의 리드는, 이들의 서열이 정규 리드 데이터세트에도 존재한다면 폐기되었다. 이어서, 매핑되지 않은 리드, 2차 정렬 및 저품질 리드는 samtools view(PMID 19505943)를 사용하여 ERE 리드 데이터세트로부터 폐기되었다. 이어서, 나머지 ERE 리드는 모든 가능한 리딩 프레임에서 ERE 폴리펩타이드로 인 실리코 번역되었다. ERE 폴리펩타이드는 정지 코돈의 위치에서 분할되었고, 하류 서열은 폐기되었고, 8개 이상의 아미노산의 상류 서열(즉, MAP의 최소 길이)만이 유지되었다. 얻어진 ERE 프로테옴은 각각의 샘플의 개인화된 정규 프로테옴과 결부되었다.

[0189] *4) mTEC+MPC k-mer 고갈에 의한 AML-특이적 프로테옴의 생성(**도 2C**).* 이 접근법을 수행하기 위하여, mTEC k-mer 고갈에 대해서 설명된 동일한 방법이 다음의 변형과 함께 사용되었다: (i) k-mer 대조군으로서 사용된 11개 MPC 샘플의 fastq 파일을 조합하는 추가의 정상 k-mer 데이터베이스가 Jellyfish로 생성되었다. 이들 샘플은 가닥 모드에서 시퀀싱되지 않았기 때문에, k-mer 데이터베이스는 -C 옵션으로 생성되었고, R1 fastq 파일은 역보완되지 않았다. (ii) mTEC 또는 MPC k-mer 데이터베이스에 존재하는 AML k-mer는 제거되었고, 그 결과 이 단계에서 필터링된 k-mer의 수는 mTEC k-mer 고갈 접근법에서보다 많았다. (iii) 정상 샘플에 의한 k-mer 고갈의 더 높은 효능 때문에, AML k-mer를 사전 필터링하기 위하여 더 낮은 발생 역치를 사용하는 것이 가능하여(**표 1** 및 **도 7C**), mTEC k-mer 고갈 단독에 비해서 이들 데이터베이스에 존재하는 k-mer의 동일성을 극적으로 변화시켰다(**도 7D**). 중요하게는, 가능한 시퀀싱 오류를 배제하기 위하여 3보다 낮은 발생 역치가 사용되지 않았다. 모든 다른 절차는 "mTEC k-mer 고갈에 의한 AML-특이적 프로테옴의 생성" 부문에서 보고된 바와 같이 수행되었다.

[0190] *5) k-mer 발현량 차이에 의한 AML-특이적 프로테옴의 생성(**도 2D**).* k-mer 차이 분석은, DE-kupl의 맞춤화된 사용, fastq 파일로부터 k-mer 데이터베이스의 생성을 수행하는 컴퓨터 파이프라인, k-mer 풍부도(abundance)의 정규화, 이들의 발생 및 이들의 샘플간 공유에 기반한 k-mer의 필터링, 통계학적 시험의 사용을 통한 두 상이한 조건에서의 샘플 간의 k-mer 풍부도 비교, 차등적으로 발현된 k-mer의 콘티그로의 조립, 콘티그의 게놈 상에의 정렬, 및 이들의 게놈 정렬에 기반한 콘티그의 주석에 기반하여 수행되었다(**도 8**)(Audoux et al., 2017). 구체적으로, DE-kupl 실행은, AML 시편을 11개의 MPC 대조군과 비교하기 위하여, 우선 다음의 파라미터 diff_method Ttest, kmer_length 33, gene_diff_method limma-voom, data_type WGS, lib_type unstranded, min_recurrence 6, min_recurrence_abundance 3, pvalue_threshold 0.05 및 log2fc_threshold 0.1을 사용하여 수행되었다. 이것은 AML 샘플과 MPC 샘플 간에 유의하게 차등적으로 발현된 33-뉴클레오타이드-길이 k-mer(FDR<0.05)의 서열 및 정규화된 계수치를 함유하고 적어도 6개 샘플(MPC 또는 AML) 중 3개의 최소 발생에 존재하는 diff-counts.tsv 파일을 리턴하였다. k-mer 필터링의 커스텀 룰이 요망되었기 때문에, DE-kupl(log2fc_threshold 0.1)에서 k-mer 배수 변화에 제한이 적용되지 않았고, diff-counts.tsv 파일에 제공된 k-mer 목록은, 모든 k-mer를 (i) 모든 MPC 샘플에서 완전히 존재하지 않도록(계수치=0)(따라서 적어도 6개의 AML 샘플에 존재); 또는 (ii) 적어도 6개의 AML 샘플(시편의 30% 초과)에 존재하고 10배 이상의 배수 변화를 갖도록; 또는 (iii) AML 샘플에서의 최저 풍부도보다도 낮은 풍부도로 단일 MPC 샘플에 존재하도록; 또는 (iv) 5배 이상의 배수 변화 및 FDR≤0.000001로 적어도 6개의 AML 샘플에 존재하도록 하기 위하여 오히려 수동으로 필터링되었다. 이들 규칙에 기반하여, 대략 41×10^6개 k-mer를 함유하는 새로운 diff-counts.tsv 파일이 생성되어, DE-kupl를 통한 k-mer 조립을 수행하는 데 사용되었고, 대략 2.1×10^6개 콘티그를 함유하는 merged-diff-counts.tsv 파일이 얻어졌다. 마지막으로, DE-kupl의 *annot* 함수를 사용하여, 생성된 콘티그를 GRCh38 인간 게놈 상에 매핑하고 주석을 달았다.

[0191] 각 AML 샘플에 대해서 개인화된 콘티그 서열을 얻기 위하여, DE-kupl annot의 DiffContigsInfos.tsv 출력을 사용하여, (적어도 2개의 k-mer의 조립으로부터 유도되는) 길이≥34개 뉴클레오타이드를 갖고, 간격, 삽입 또는 결실 없이 정렬된 모든 콘티그의 bed 파일을 구축하였다(N/D/I 없는 CIGAR). 다음에, 이 bed 파일 및 bedtool, samtool 및 bcftool 슈트를 사용하여, 각 AML 샘플(samtools mpileup -C50 -uf ref_genome.fasta sample.bam | bcftools call -c | vcfutils.pl vcf2fq -d 8 -D 100 | awk '/^@chr.$|^chr..$|^@GL.....$|^@KI.....$/,/^+$/' | sed '/^+/d' | tr "@" ">" > consensus.fasta)의 bam 파일(리드는 별표를 사용하여 GRCh38에 매핑됨, "*개인화된 정규 프로테옴의 생성" 부문 참조*)로부터 생성된

공통 게놈으로부터 개인화된 콘티그 서열(bedtools getfasta -fi consensus.fasta -bed contigs.bed -name >> output.fasta)을 추출하였다.

[0192] 리드(N)에 의해 포함되지 않은 콘티그의 부분은 sed(sed -E "s/NNN+/\n/g")로 제거되었고, 모든 콘티그는 fasta 파일에 기록되었다. 간격, 삽입 또는 결실(그리고 공통 게놈으로부터 검색될 수 없음)과 정렬되고 DiffContigsInfos.tsv에서 관련 샘플에 의해 표현된 바와 같이 보고된 콘티그의 서열이 이 fasta 파일에 추가되었다. 마지막으로, (이전에 공개되었거나 또는 pyGeno(Daouda et al., 2016; Laumont et al., 2018)에 포함되어 있는) 사내 파이썬 스크립트를 사용함으로써, 콘티그는 6-프레임 번역되었고, (단일 염기 돌연변이와 중첩하는 콘티그가 다수의 상이한 아미노산 서열을 코딩할 수 있기 때문에) 모호한 아미노산 서열은 모든 가능한 서열로 변형되었고, 아미노산 서열은 내부 정지 코돈에서 분할되었고, 얻어진 하위서열은 각 샘플의 각각의 개인화된 정규 프로테옴과 결부되었다.

[0193] *6) 데이터베이스 크기의 검증 - **도 9B 내지 C.*** 이 연구에 사용된 4가지 프로테오게노믹 접근법에 사용된 MS 데이터베이스가 정규(개인화된) 프로테옴 데이터베이스와 비교해서 다양하게 부풀려진 크기를 나타낸 것을 고려할 때, 이러한 더 큰 크기가 MS 식별에 미치는 영향을 조사하였다. 우선, 각 접근법으로 식별된 펩타이드의 누적 수를 19개 AML 샘플에 걸쳐서 비교하였다(**도 9B**). 이것은 접근법들 간의 데이터베이스 크기의 상당한 차이에도 불구하고, 식별된 펩타이드의 수가 정규 프로테옴과 비교해서 단지 약간만(ERE 접근법의 경우 최대 대략 9%) 달라진 것을 나타내었다. 다음에, 각 데이터베이스가 각 샘플의 각각의 정규 개인화된 프로테옴과 결부되었기 때문에, 적절한 크기의 데이터베이스가 정규 프로테옴 단독에 대한 유사한 동일성의 정규 단백질-유래 펩타이드의 식별을 허용해야 한다는 것이 추론되었다. **도 9C**에 도시된 바와 같이, 모든 AML 샘플에 걸쳐서 각 접근법에서 식별된 단백질-코딩으로 주석달린 펩타이드의 대다수(88.2% 내지 96.2%)가 정규 프로테옴 단독에 기초하여 식별된 것과 공통적이었다. 이러한 관찰에 기초하여, 다양한 데이터베이스의 크기가 신뢰할 수 있는 MS 식별에 적합하다는 결론이 내려졌다.

[0194] *MHC-연관 펩타이드의 단리*

[0195] W6/32 항체(BioXcell)를, 슬리리 1㎖당 항체 1㎎의 비율로 PureProteome protein A 자기 비드(Millipore)와 함께 실온에서 60분 동안 PBS에서 인큐베이션하였다. 항체를 기재된 바와 같이 다이메틸피멜리데이트를 사용하여 자기 비드에 공유적으로 가교결합시켰다. 비드를 PBS pH 7.2 및 0.02% NaN3에서 4℃에 보관하였다. 동결된 세포 펠릿 샘플(9천8백만개 세포/펠릿)의 경우, 세포를 해동시키고, 1㎖의 PBS pH 7.2에 재현탁시키고, PBS pH 7.2, 프로테아제 저해제 칵테일(Protease inhibitor cocktail)(Sigma)이 보충된 1% (w/v) CHAPS(Sigma)를 함유하는 1㎖의 세제 완충제를 첨가함으로써 가용화시켰다. 종양 샘플의 경우, 조직을 작은 조각(입방체, 대략 3㎜ 크기)으로 절단하고 프로테아제 저해제 칵테일을 함유하는 빙랭 PBS 5㎖를 첨가하였다. 조직 조각을 먼저 속도 20000 rpm으로 설정된 Ultra Turrax T25 균질화기(IKA- Labortechnik)를 사용하여 20초 동안, 이어서 25000 rpm으로 설정된 Ultra Turrax T25 균질화기(IKA- Labortechnik)를 사용하여 20초 동안 2회 균질화시켰다. 이어서, 550㎕의 빙랭 10X 용해 완충제(5% w/v CHAPS)를 샘플에 첨가하였다. 세포 펠릿 및 종양 샘플을 4℃에서 텀블링하면서 60분 인큐베이션하고, 이어서 4℃에서 20분 동안 10000g에서 회전시켰다. 상청액을 샘플당 1㎎의 W6/32 항체 공유-가교결합된 protein A 자기 비드를 함유하는 새로운 튜브로 옮기고, 4℃에서 180분 동안 텀블링하면서 인큐베이션하였다. 샘플을 자석 위에 놓고 결합된 MHC I 복합체를 자기 비드로 회수하였다. 자기 비드를 우선 8×1㎖ PBS로, 이어서 1×1㎖의 0.1X PBS로, 마지막으로 1×1㎖의 물로 세척하였다. MHC I 복합체는 0.2% 폼산(FA)을 이용한 산성 처리에 의해 자기 비드로부터 용출시켰다. 임의의 잔류 자성 비드를 제거하기 위하여, 용출액을 2.0㎖ Costar mL Spin-X 원심관 필터(0.45㎛, Corning)로 옮기고 855g에서 2분 동안 회전시켰다. 펩타이드를 함유하는 여과액은, 20개의 1㎜ 직경 옥타데실(C-18) 고체상 추출 디스크(EMPORE)가 패킹된 자체 제작된 스테이지 팁을 이용해서 MHC I 소단위(HLA 분자 및 β-2 매크로글로불린)으로부터 분리시켰다. 스테이지 팁은 먼저 메탄올로, 이어서 0.2% 트라이플루오로아세트산(TFA) 중 80% 아세토나이트릴(ACN)로, 마지막으로 0.2% FA로 사전 세척하였다. 샘플을 스테이지 팁에 로딩하고 0.2% FA로 세척하였다. 펩타이드를 0.1%TFA 중 30% ACN으로 용리시키고, 진공 원심분리를 사용하여 건조시키고, 이어서 MS 분석될 때까지 -20℃에서 보관하였다.

[0196] *질량분광분석*

[0197] 건조된 펩타이드 추출물을n 4% 폼산에 재현탁시키고, EasynLC II 시스템에서 600 nL/분의 유량 및 0%에서 30%까지의 아세트나이트릴(0.2% 폼산)의 56분 구배(10H005) 또는 106분 구배(모든 다른 샘플)로 자체 제작된 C18 분석 칼럼(15㎝×150㎛ i.d., C18 Jupiter Phenomenex로 패킹됨)에 로딩하였다. 샘플은 1.6kV에서 Nanospray 2

소스를 사용하여 양이온 모드에서 Q-Exactive HF 질량 분석계(Thermo Fisher Scientific)로 분석하였다. 60,000 분해능으로 획득된 각각의 전체 MS 스펙트럼에 이어서 20 MS/MS 스펙트럼이 후속되는데, 여기서 30,000의 분해능, 5×10^4(10H005) 또는 2×10^4(모든 다른 샘플)의 자동 이득 제어 목표, 100ms(10H005) 또는 500ms(15H023, 15H063, 15H080, 05H149) 또는 800ms(모든 다른 샘플)의 주입 시간 및 25%의 충돌 에너지로 MS/MS 시퀀싱을 위하여 가장 풍부한 다중 하전 이온이 선택되었다.

[0198] *합성 펩타이드*

[0199] 충분한 양의 물질이 이용 가능할 때, TSA^{hi}의 아미노산 서열은, (Zhao *et al.*, Cancer Immunol Res. 2020 Feb 11 doi: 10.1158/2326-6066.CIR-19-0541. [Epub ahead of print])에 이미 기재된 바와 같이, 합성 펩타이드로 더욱 검증되었다.

[0200] *생물정보 분석*

[0201] 모든 분석은 트리밍된 데이터에 대해 수행되었고, 모든 정렬은 달리 언급되지 않는 한 앞서의 부문에 기재된 바와 같이 별표로 이루어졌고, 모든 정렬은 GRCh38.88에서 이루어졌다.

[0202] 모든 액체 크로마토그래피(LC)-MS/MS (LC-MS/MS) 데이터는 PEAKS X(Bioinformatics Solution Inc.)를 사용하여 관련 데이터베이스에 대해 검색되었다. 펩타이드 식별을 위하여, 전구체 및 단편 이온에 대해 허용 오차는 각각 10ppm 및 0.01Da로 설정되었다. 산화(M) 및 탈아마이드화(NQ)의 발생은 가변 변형으로 설정되었다.

[0203] *1) MAP의 식별.* 펩타이드 식별 후, 고유한 펩타이드의 목록이 각 샘플에 대해 얻어졌고, 5%의 거짓 발견율(false discovery rate: FDR)이 펩타이드 스코어에 적용되었다. 샘플의 HLA 대립유전자에 대한 결합 친화도는 NetMHC 4.0(Andreatta and Nielsen, 2016)으로 예측되었고, 백분위수 순위가 2% 이하인 8 내지 11개-아미노-산-길이의 펩티드만 추가의 주석에 사용되었다.

[0204] 2) *관심 MAP(MOI)의 식별 및 검증.* 두 k-mer 고갈 접근법의 경우, 이것은 앞서 기재된 것(Laumont et al., 2018)과 유사한 접근법을 사용하여 수행되었다. 간단히 말해서, 각 MAP 및 이의 코딩 서열은 각각 관련 AML 및 정상의 정규 프로테옴(위에서 상세히 설명된 바와 같이, 모든 mTEC 및 MPC에 대해서 구축됨) 또는 암 및 정상의 24개-뉴클레오티드-길이 k-mer 데이터베이스(위에서 상세히 설명된 바와 같이, 조합된 mTEC 또는 조합된 MPC로부터 구축됨)에 대해서 쿼리되었다(queried). 정상의 정규 프로테옴에서 검출된 MAP는 그의 코딩 서열 검출 상태에 관계없이 배제되었다. 정상의 정규 프로테옴에서도 정상의 k-mer에서도 검출되지 않은 MAP가 MOI로 플래깅되었다. 두 k-mer 데이터베이스에는 존재하지만 두 정규 프로테옴에 존재하지 않는 MAP는, MOI로 플래깅되기 위하여, 정상의 샘플과 비교해서 AML에서 적어도 10배만큼 과발현된 그들의 RNA 코딩 서열을 가질 필요가 있었다. 마지막으로, 몇 가지 RNA 서열(상이한 단백질로부터 유래)에 대응하는 MAP는, 그들의 각각의 코딩 서열이 일관되게 이들을 MOI로 플래깅한 경우에만 MOI로 플래깅될 수 있었다.

[0205] ERE 접근법의 경우, "Yes", "Maybe" 또는 "No"의 ERE 상태가 ERE 및 개인화된 정규 프로테옴에서 그의 아미노산 서열의 존재에 기반하여 각각 개별 MAP에 부여되었다. "Maybe" 후보물질의 경우, ERE 리드 및 정규 리드 데이터세트에서 펩타이드의 코딩 서열(즉, 펩타이드의 24-뉴클레오타이드-길이 k-mer 세트의 최소 발생)의 발현 수준이 컴퓨팅되었다. ERE 리드 데이터세트에서보다 적어도 10배 높은 발현을 갖는 "Maybe" 후보물질만이 ERE MAP로 간주된다. 이어서, 나머지 ERE MAP 후보물질은, 펩타이드의 코딩 서열이 생식계열 다형성을 함유하고 ERE 서열 및 정규 주석달린 서열(적용 가능한 경우)과 비교된 적절한 배향을 갖는지를 결정하기 위하여 IGV(Robinson et al., *Nat Biotechnol*. 2011 Jan;29(1):24-6)에서 수동으로 검증되었다.

[0206] 차등 k-mer 접근법의 경우, MAP의 전체 목록은 MOI 후보물질로 플래깅될 AML-특이적 프로테옴에 쿼리되었다. 다음에, (다음 부문에 기재된 절차 후에) 각 MAP의 RNA 발현은 DE-kupl에서 대조군으로서 사용된 11개 MPC에서 그리고 19개 AML 시편에서 평가되었고, 정상 샘플과 암 샘플 사이에서 5의 최소 배수 변화를 갖는 모든 MAP를 MOI로 플래깅되었다. MAP RNA-발현 평가 절차가 정량화를 수행하기 위하여 참조 게놈을 기반으로 하기 때문에, 돌연변이로부터 유래된 후보물질 MOI는 적절하게 정량화될 수 없었고, MOI 후보물질로서 체계적으로 플래깅되었다. 식별된 각 AML 샘플에서 각 MOI의 RNA 수준에서의 존재를 명확하게 검증하기 위하여, MOI 코딩 서열은 DE-kupl의 DiffContigsInfos.tsv 출력으로부터 검색되었고, 관련 fastq 파일(정방향 R2 fastq에서의 서열 및 역방향 R1 fastq에서의 역보완)에 쿼리되었다. 이 시험에 통과하지 못한 MOI는 폐기되었다.

[0207] 모든 MOI 후보물질 목록(4가지 다른 접근법)의 경우, 류신 및 아이소류신 변이체는 표준 MS 접근법으로 구별할 수 없기 때문에, 각각의 목록을 검사하고, 기존의 변이체가 비-MOI로 플래깅된 MOI는 변이체보다 더 높은 RNA

발현을 제시하지 않는 한 폐기되었다. 모든 MOI의 MS/MS 스펙트럼은 임의의 가짜 식별을 제거하기 위하여 수동으로 검사되었다. 마지막으로, 게놈 위치는 BLAT(UCSC 게놈 브라우저로부터의 툴)를 사용하여 참조 게놈에 코딩 서열을 함유하는 리드를 매핑함으로써 모든 MOI에 할당하였다. 리드가 일치하는 게놈 위치와 정합(match)하지 않거나 또는 초가변 영역(예컨대, MHC, Ig 또는 TCR 유전자)과 정합하는 MOI는 배제되었다. 일치하는 게놈 위치를 갖는 것의 경우, IGV를 사용하여 공지된 생식계열 다형성(dbSNP149)과 중첩하는 코딩 서열을 갖는 MOI를 배제하였다.

[0208] *3) RNA-Seq 데이터에서 MAP 코딩 서열의 정량화.* 각 MAP의 RNA 발현을 명확하게 평가하기 위하여, 모든 MAP 아미노산 서열은 모든 가능한 뉴클레오타이드 서열로 역번역되었다. 다음에, 이들 모든 가능한 서열은 -n 1000000 옵션과 함께 GSNAP(Wu et al., 2016)으로 게놈 상에 매핑되어, 주어진 MAP를 코딩할 수 있는 모든 게놈 영역을 찾았다. 스플라이스 부위가 중첩되는 서열에 의해 코딩된 MAP를 확실하게 캡처하기 위하여, 가능한 MAP 코딩 서열도 전사체(cDNA 및 비-코딩 RNA)에 매핑되어, 참조 전사체 서열의 많은 부분(80개 뉴클레오타이드)을 추출하고(--길이 80 옵션을 가진 samtools faidx) 이어서 참조 게놈(GSNAP, --use-splicing 및 --novelsplicing=1 옵션을 가짐)에 매핑되었다. 상이한 TSA 발견 파이프라인에 의해 생성된 MOI의 경우, 코딩 서열을 함유하는 모든 리드의 게놈 정렬이 또한 수행되었다. GSNAP의 출력은 주어진 MAP을 코딩하기 쉬운 모든 가능한 게놈 영역을 함유하는 bed 파일을 생성하기 위하여 서열과 참조 서열 간에 완전한 정합만 유지하도록 필터링되었다. samtools view(-F256 옵션), grep 및 wc(-l 옵션)을 사용함으로써, 각각의 목적하는 RNA-Seq 샘플(예컨대, AML, GTEX 또는 정상 샘플)에서 각각의 게놈 위치에 MAP 코딩 서열을 함유하는 리드의 수가 별표(bam 파일)를 가진 참조 게놈에 정렬되어 계수되었다. 마지막으로, 주어진 MAP에 대한 (상이한 영역 및 코딩 서열로부터의) 모든 리드 계수치가 합산되고, 각 평가된 샘플에서 시퀀싱된 리드의 총수에 정규화되어, 백만당 리드(reads-per-hundred-million: RPHM) 계수치를 얻었다.

[0209] *4) 면역원성 평가.* MOI의 면역원성 예측은 Repitope(Ogishi and Yotsuyanagi, 2019)로 수행되었다. 특징 계산(Feature computation)은 미리 정의된 MHCI_Human_MinimumFeatureSet 변수로 수행되었고, 패키지의 Mendeley 저장소(https://data.mendeley.com/datasets/sydw5xnxpt/1)에 제공된 FeatureDF_MHCI 및 FragmentLibrary 파일을 업데이트하였다(2019년 7월 12일).

[0210] *5) AML 환자에 의한 MOI 제시 및 발현.* 주어진 MAP(무차별 결합제)를 제시 가능한 모든 가능한 HLA 대립유전자를 식별하기 위하여, MHC클러스터 온라인 툴(http://www.cbs.dtu.dk/services/MHCcluster/)(Thomsen et al., 2013)이 사용되었다. 클러스터링 값이 0.4 이하인 HLA 대립유전자가 동일한 MAP를 제시 가능한 것으로 간주되었다. Leucegene 코호트에서 AML 환자에 의한 MOI 제시를 평가하기 위하여, 이들의 HLA 유형은 우선 Optitype으로 결정되었고, 주어진 MOI는 RNA 수준에서의 발현이 2 rphm 초과인 경우(0 rphm 대신, 제시 가능성을 최대화하기 위하여) 그리고 환자가 MOI를 제시 가능한 HLA 대립유전자를 발현하는 경우(각 개발된 MOI의 제시 분자의 원래의 식별을 위한 NetMHC4.0 및 무차별 결합제의 식별을 위한 MHC클러스터에 의해 제시된 바와 같이) 제시된 것으로 간주되었다. 환자가 동일한 MOI를 제시 가능한 두 상이한 HLA 대립유전자를 발현하는 경우, MOI는 두번 제시된 것으로 간주되었다.

[0211] 높은 TSA^{high} 발현과 관련된 분자 특징을 평가하기 위하여, TSA^{high}는, 이 환자에서의 발현이 전체 코호트에 걸쳐서 (null이 아닌 값에만 기반하여 계산된) 중앙 발현보다 높은 경우 주어진 환자에서 발현된 것으로 간주되었다. 고도로 발현된 TSA^{high}의 총수(# HE-TSA^{hi})가 각 환자에 대해서 계산되었고, 유전자 발현과의 상관관계 분석 및 돌연변이 또는 기타 임상 특징과의 연관을 수행하기 위하여 사용되었다(다음 부문 참조).

[0212] *6) 생존율 분석.* Leucegene 코호트의 374명의 환자에 대한 생존율 데이터는 Leucegene 팀(https://leucegene.ca)으로부터의 친절한 선물이었다. 생존율 분석은, 위에서 기재된 바와 같이 계산된 높은 수의 HLA-TSA^{hi} 복합체(TSA^{hi}의 HLA-제한된 제시)와 임상 결과(전체 생존율)와의 연관성을 평가하기 위하여 수행되었다. 환자는 이들이 제공할 수 있었던 TSA^{hi}의 총수에 따라서 두 군으로 분리되었다: 높은 발현자HLA-TSA^{hi} 계수치의 상위 사분위수) 및 낮은 발현자(기타 모든 환자). 생존율은 Kaplan-Meier 곡선을 사용하여 두 군 간에 비교했으며, 유의성은 GraphPad Prism v7.0에서 로그 순위법으로 평가하였다. R에서의 패키지 생존율 분석 v0.1.1로 수행된 다변량 분석에서, 연령은 연속 변수로서 통합되었고, 돌연변이는 존재/부재(1/0)로서 코딩되었고, 세포유전학 위험의 평가는 개별 군으로서 처리되었는데, 중간 대 유리한 위험과 불리 대 유리한 위험에 대해서 행해졌다.

[0213] *7) 돌연변이 분석.* NPM1, FLT3-ITD, FLT3-TKD, IDH1(R132) 및 이중대립유전자 CEBPA에 대한 돌연변이 데이터는

Leucegene 코호트에 대한 이전에 공개된 데이터(Audemard et al., 2019; Lavallee et al., 2016)로부터 검색되었다. ASXL1, TP53, DNMT3A, IDH2(단지 R140 및 R172), WT1, RUNX1 및 TET2의 돌연변이는 Freebayes로 검출되었고, 하기의 돌연변이를 제거하기 위하여 필터링되었다: (i) 변이체 대립유전자 빈도(VAF) < 20%를 갖는 것; (ii) COSMIC 데이터베이스(https://cancer.sanger.ac.uk/cosmic)에서 SNP로서 플래깅되는 것; (iii) 낮은 추정 충격(5'UTR 조기 개시 코돈 이득 변이체; 스플라이스 영역 변이체 및 동의어 변이체(synonymous variant); 정지 보유 변이체; 동의어 변이체)을 갖는 것; (iv) FATHMM-XF(http://fathmm.biocompute.org.uk/fathmm-xf/)(Rogers et al., 2018)에 의해 예측된 바와 같이 단백질 구조 및 기능에 양성 영향을 미치는 미스센스 SNP; (iv) AAAAA+ 또는 TTTTT+와 연루된 삽입 및 결실; (v) COSMIC 데이터베이스(Tate et al., 2018)에서 생식 계열로서만 플래깅된 돌연변이.

[0214] *8) 유전자 발현 분석.* 모든 전사체 발현 정량화는 디폴트 파라미터와 함께 kallisto v0.43.0으로 수행되었다. Kallisto의 전사체-수준 계수 추정치는 R 패키지 tximport를 사용하여 유전자-수준 계수치로 변환되었다. EdgeR은 TMM 알고리즘을 사용하여 계수치를 정규화하고 백만당 계수치(count-per-million: cpm) 값을 출력하는 데 사용되었다. 단백질-코딩 유전자(앙상블(Ensembl)의 BioMart 툴에서 보고된 바와 같음, useast.ensembl.org/biomart)만 추가의 분석을 위하여 유지되었다. 각 유전자 발현과 HE-TSA^{hi} 계수치 간의 체계적인 피어슨 상관관계는 R에서의 cor.test 함수로 수행되었다. 상관관계는 모든 환자, 비-NPM1/FLT3-ITD/DNMT3A 돌연변이된 환자 및 비-FAB-M1 환자에 대해서 수행되었다. p-값은 Benjamini 및 Hochberg 방법(R에서의 p.adjust)과 다중 비교를 위하여 보정되었고, 3개의 상관관계 분석 중 적어도 하나에서 FDR < 0.00001, 3가지 분석에서의 FDR < 0.001, 3가지 분석에서 일관된 상관관계 계수(양 또는 음) 및 적어도 하나의 분석에서 상관관계 계수 > 0.3 또는 < -0.3을 갖는 유전자만 하류 처리를 위하여 유지되었다.

[0215] t-분산 확률 이웃 임베딩(t-SNE) 분석은 tximport(tpm ≥1인 경우 발현 = 1 그리고 tpm < 1인 경우 발현 = 0)에 의해 Kallisto의 전사체-수준 풍부도 추정치를 유전자 풍부도 추정치로 종합한 것으로부터 얻어진 발현된 유전자의 동일성에 대해 Rtsne 패키지로 수행되었다. 이 분석에는 단백질-코딩 유전자(MAP를 생성하기 가장 쉬움)만 사용되었다.

[0216] *9) GO 용어 및 농축 맵 분석.* 생물학적-과정 유전자-온톨로지(GO) 용어 과다표시는, Cytoscape v3.7.2에서 BiNGO v3.0.3(Maere et al., 2005)을 사용하고, 초기하학적 시험을 사용하고, 그리고 0.005 이하의 FDR-조정된 P값의 유의성 컷오프를 적용하여 수행되었다. BiNGO로부터의 출력은 Cytoscape에서의 EnrichmentMap v3.2.1(Merico et al., 2010)에 임포트하여, 중복 GO 용어를 클러스터링하고 그 결과를 시각화하하였다. EnrichmentMap은 0.25의 Jaccard 유사성 계수 컷오프, 0.001의 P 값 컷오프 및 0.005의 FDR-조정된 컷오프를 사용하여 생성되었다. 네트워크는디폴트 세팅과 600회 반복으로 Cytoscape에서의 디폴트 "Prefuse Force-Directed Layout"을 사용하여 시각화되었다. 유사한 GO 용어의 군은 수동으로 동그라미를 쳤다.

[0217] *10) 인트론 보유 및 NMF 클러스터링.* 전체 Leucegene 코호트 및 11개의 주된 MPC 샘플의 인트론 보유(IR) 분석은 IRFinder v1.2.5(Middleton et al., 2017)로 수행되었다. IRatio≥10%을 갖는 인트론(전사체의 10% 이상에서 보유된 인트론) 및 최소 3개의 리드 범위는 보유된 것으로 간주되었다. 인트론은 적어도 2개의 AML 샘플에서 보유되고 임의의 MPC 샘플에서 보유되지 않은 것들만 유지시키도록 필터링되었다. (전체 코호트에 걸쳐서 IRatio의 변동 계수에 의한) 10% 가장 가변적인 인트론(6988개 인트론)이 추가의 분석을 위하여 선택되었다. 무감독 공통 클러스터링 결과는, 선택된 인트론의 IRatio에 대한 R에서의 NMF v0.21.0(Gaujoux and Seoighe, 2010) 패키지로, 디폴트 Brunet 알고리즘, 및 순위 조사 및 클러스터링 실행에 대한 200회 반복으로 생성되었다. 클러스터 결과는, 3 내지 15개의 클러스터를 갖는 클러스터링 솔루션에 대해서, 컨센서스 멤버십 매트릭스(consensus membership matrix)의 공표형 스코어 및 평균 실루엣 폭의 프로파일을 고려함으로써 선택되었다.

[0218] 풍부도 히트맵은 NMF 메타유전자(metagene)(W 매트릭스) 출력 파일에서 상위 2% 인트론을 식별함으로써 생성되었다. 중복된 이름을 제거하면 1211개의 인트론의 목록이 생성되었다. 이들 인트론 IRatios의 행렬은 각 Leucegene 샘플에 대해 생성되었고, NMF 클러스터링 출력과 정합시키도록 재정렬되었고, R에서의 heatmap.3 패키지를 사용하여, 센터링된 상관관계 거리 메트릭 및 완전한 연쇄로 인트론의 계측정 클러스터링을 수행하였다.

[0219] *ELISPOT 검정*

[0220] 1) *단핵구-유래 수지상 세포의 생성.* 단핵구-유래 수지상 세포를, 앞서 기재된 바와 같이(Vincent *et al.*, *Biology of Blood and Marrow Transplantation: Journal of the American Society for Blood and Marrow Transplantation*, 22 Oct 2013, 20(1):37-45; Laumont *et al.*, *Nat Commun.* 2016 Jan 5;7:10238), 냉동된

PBMC로부터 생성하였다. 간단히 말해서, DC는 5% 인간 혈청(Sigma-Aldrich), 피루브산나트륨(1mM), IL-4(100 ng/㎖, Peprotech) 및 GM-CSF(100 ng/㎖, Peprotech)가 보완된 X-VIVO™ 15 배지(Lonza Bioscience)에서 8일 동안 배양에 의해 부착 PBMC 분획으로부터 준비하였다. 배양 7일 후, DC를 IFN-γ(1000 IU/㎖, Gibco) 및 LPS(100 ng/㎖, Sigma Aldrich)로 하룻밤 성숙시켰다. DC에 성숙 과정 후 2시간 동안 2 ㎍/㎖의 펩타이드를 로딩하고, 이어서 방사선 조사하였다(T-DC 배양물에서 APC로서 사용되기 전에 40Gy). 대조군의 경우, DC를 MelanA, NS3 및 Gag-A2 펩타이드(3가지 모두 HLA-A*02:01에 결합됨)를 함유하는 혼합물로 펄싱하였다.

[0221] 2) *시험관내 펩타이드-특이적 T 세포 확장.* 해동된 PBMC는 우선 인간 $CD8^+$ T 세포 단리 키트(Miltenyi Biotech)를 사용하여 농축된 $CD8^+$ T-세포였고, 1:10의 APC:T 세포비로 자가 펩타이드-펄싱된 DC와 공동 인큐베이션되었다. 확장되는 T 세포를 8% 인간 혈청(Sigma-Aldrich), L-글루타민(Gibco) 및 사이토카인이 보충된 고급(Advanced) RPMI 배지(Gibco)에서 (7일마다 펄싱된-DC 재자극으로) 4주 동안 배양하였다. 첫 번째 공동 배양 주 동안, 배지에 IL-12(10 ng/㎖) 및 IL-21(30 ng/㎖)을 첨가하였다. 2일 후에, IL-2(100 UI/㎖)를 사이토카인 믹스에 첨가하였다. 두 번째 주에, 배지에 IL-2(100 UI/㎖), IL-7(10 ng/㎖), IL-15(5 ng/㎖) 및 IL-21(30 ng/㎖)을 첨가하였다. 공동배양의 마지막 2주 동안, IL-2(100 UI/㎖), IL-7(10 ng/㎖) 및 IL-15(5 ng/㎖)를 사용하였다. 적절한 사이토카인 믹스가 보충된 배지를 2일마다 공동배양물에 첨가하였다. 공동 배양 4주의 말기에, ELISPOT 검정을 수행하기 위하여 세포를 수거하였다.

[0222] 3) *IFNγ ELISPOT 검정.* ELISpot 인간 IFNγ(R&D Systems, 미국 소재) 키트는 실험을 수행하기 위해 제조업체의 권장 사항에 따라서 사용하였다. 이어서, 수거한 $CD8^+$ T 세포를 플레이팅하고, 자극제 세포로서 사용된 광조사된 펩타이드-펄싱된 PBMC(40Gy)의 존재 하에 37℃에서 24시간 동안 인큐베이션시켰다. 음성 대조군으로서, 분류된 $CD8^+$ T 세포를 조사된 비펄싱 PBMC와 함께 인큐베이션시켰다. 반점이 제조사 프로토콜에 언급된 바와 같이 나타났고, ImmunoSpot S5 UV 분석기(Cellular Technology Ltd, 오하이오주 셰이커하이츠 소재)를 사용하여 계수되었다. IFN-γ 생산은 음성 대조군 웰로부터 반점 계수치를 뺀 후 10^6개 $CD8^+$ T 세포당 펩타이드-특이적 반점-형성 세포(SFC)의 수로서 표현되었다.

[0223] *면역원성 예측*

[0224] MOI의 면역원성 예측은 Repitope(Ogishi and Yotsuyanagi, 2019)로 수행되었다. 특징 계산은 사전 정의된 MHCI_Human_MinimumFeatureSet 변수로 수행되었고, 패키지의 Mendeley 저장소(https://data.mendeley.com/datasets/sydw5xnxpt/1)에 제공된 FeatureDF_MHCI 및 FragmentLibrary 파일을 업데이트하였다(2019년 7월 12일).

[0225] *TCR 및 세포독성 T 세포 시그니처 분석*

[0226] TCR 레퍼토리 분석은 TRUST4 소프트웨어(Li et al., 2017) 및 디폴트 파라미터와 함께 437명의 Leucegene 환자의 RNA-seq 데이터에 대해 수행되었다. T 세포의 클론형 다양성은 킬로 TCR 리드(CPK)당 TCR CDR3(완전 및 부분)의 수를 정규화함으로써 추정되었다. MOI와 TRUST4에 의해 검출된 완전한 TCR베타 CDR3 아미노산 서열 간의 상호작용의 ERGO(Springer et al., 2020) 예측은 훈련용 데이터베이스로서 VDJdb 및 Autoencoder 기반 모델을 사용하여 무료로 이용 가능한 웹포탈(http://tcr.cs.biu.ac.il/)을 통해서 이루어졌다.

[0227] 세포독성 T 세포 시그니처 분석을 위하여, 환자당 예측된 HLA-TSA^{hi} 쌍의 계수치는 rphm 발현이 2 이상인 TSA^{hi}의 계수치로 나누어서 정규화된 TSA^{hi} 제시 수준을 얻었다. HLA-TSA^{hi} 계수치가 없거나 또는 진단 시 수집되지 않은 환자 샘플은 분석으로부터 폐기되었다. 나머지 361명의 환자는 그들의 정규화된 TSA^{hi} 제시 수준에 따라서 그룹화하고, 분포의 중앙값보다 높은 환자는 차등적 유전자 발현 분석을 통해서 다른 것(중앙값 미만)과 비교되었다. 이 분석은 R3.6.1에서 수행되었다. 원시 리드 계수치는 백만 계수치(cpm)로 변환되었고, 라이브러리 크기에 대해서 정규화되었고, edgeR 3.26.8(Robinson et al., 2010) 및 limma 3.40.6(Ritchie et al., 2015)을 사용하여 적어도 2개의 샘플에서 cpm>1인 유전자를 유지시킴으로써 낮게 발현된 유전자를 필터링하였다. 이것에 이어서, limma의 lmfit를 사용한 선형 모델링 및 voom 형질전환이 행해졌다. 마지막으로, 조정된 t-통계(moderated t-statistics)는 eBayes로 계산되었다. p-값≤0.01 및 $-0.3 \geq \log_2(FC) \geq 0.3$인 유전자는 유의하게 차등적으로 발현된 것으로 간주되었다.

[0228] *사이토카인 분비 검정 및 덱스트라머*

[0229] (Janelle et al., 2015)에 기초한 펩타이드 로딩된 단핵구-유래 수지상 세포 및 사이토카인을 사용한 3회 자극 후, 1,0×10^6개 세포를 다이메틸 설폭사이드(DMSO), 5 ㎍/㎖의 관심 펩타이드, 5 ug/㎖의 대조 펩타이드(음성 대조군) 또는 50 ng/㎖의 포볼(phorbol) 12-미리스테이트 13-아세테이트(PMA) 및 500 ng/㎖ of 이오노마이신(ionomycin)(양성 대조군, Sigma-Aldrich)과 함께 7.5 ㎍/㎖의 Brefeldin A(Sigma-Aldrich, 온타리오주 오크빌 소재)의 존재 하에 4시간 동안 인큐베이션시켰다. 이어서, 세포를 세포 표면 항체로 염색하고, 제조사의 지침(BD Biosciences, 온타리오주 미시소거 소재)에 따라서 세포내 염색을 위해 Cytofix/Cytoperm 완충제를 사용하여 고정 및 투과시켰다. 투과된 세포를 IFNɣ, IL-2 및 TNFa(BD Biosciences)에 대해서 지향된 항체와 함께 4℃에서 20분 동안 인큐베이션시키고, 획득 전에 2% 소태아 혈청(FBS; ThermoFisher, 미국 매세추세츠주 월섬 소재)가 보충된 인산염 완충 식염수(PBS)에 재현탁시켰다. 획득은 LSRII 유세포분석기(BD Biosciences)로 수행되었고, 데이터는 FlowJo™ V10 소프트웨어(BD Biosciences)를 사용하여 분석되었다. 다량체 염색을 위하여, 1,0×10^6개의 세포를 맞춤 제작된 형광 덱스트라머(Immudex, 덴마크 코펜하겐 소재)로 4℃에서 45분 동안 염색하고, 이어서 CD8 단클론성 항체(eBiosciences, 캘리포니아주 샌디에고 소재)로 4℃에서 30분 동안 염색하였다. LSRII 유세포분석기(BD Biosciences)로 획득 전에 세포를 PBS 2% FBS로 세척하였다. 데이터는 FlowJo™ V10 소프트웨어(BD Biosciences)를 사용하여 분석되었다.

[0230] *FEST 검정*

[0231] FEST 검정의 경우, T 세포는 앞서 설명된 바와 같이 약간의 변형(Danilova et al., 2018)으로 배양되었다. 간단히 말해서, 0일째에, 건강한 공여체(BioIVT)로부터 해동된 PBMC는 인간 팬(Human Pan) T-세포 단리 키트(Miltenyi)를 사용하여 농축된 T-세포였다. T 세포를 50 ㎍/㎖ 젠타마이신(ThermoFisher Scientific) 및 1% Hepes가 보충된 AIM V 배지에서 2×10^6/㎖로 재현탁시켰다. T 세포-음성 분획을 30Gy에서 조사하고, 세척하고 50 ㎍/㎖ 젠타마이신 및 1% Hepes가 보충된 AIM V 배지에서 2.0×10^6/㎖로 재현탁시켰다. T 세포 및 조사된 T 세포-고갈된 세포 둘 다의 웰당 1㎖를 3개의 TSA^{hi} 풀(5개 TSA^{hi}/풀, 각 TSA에 대해서 1μM 최종 농도) 중 하나와 또는 펩타이드 없이 12-웰 플레이트에 첨가하였다. 세포를 37℃, 5% CO2에서 10일 동안 배양하였다. 3일째 및 7일째에, 배양 배지의 절반을, 100 IU/㎖ IL-2, 50 ng/㎖ IL-7 및 50 ng/㎖ IL-15(3일) 및 200 IU/㎖ IL-2, 50 ng/㎖ IL-7 및 50 ng/㎖ IL-15(7일)를 함유하는 신선한 배양 배지로 교체하였다. 10일째에, 세포를 수거하고, $CD8^+$ 세포를 인간 $CD8^+$ T 세포 단리 키트(Miltenyi)를 사용하여 더욱 단리시켰다. 음성 대조군으로서, $CD8^+$ T 세포를 또한 동일한 건강한 공여체의 신선하게 해동된 배양되지 않은 PBMC로부터 단리시켰다. Qiagen DNA 혈액 미니 키트(Qiagen)를 사용하여 $CD8^+$ T 세포로부터 DNA를 추출하였다. TCR Vβ CDR3 시퀀싱은 ImmunoSEQ 플랫폼(Adaptive Biotechnologies)의 조사 솔루션(survey resolution)을 사용하여 수행되었다. immunoSEQ 포탈로부터 엑스포트된 원시 데이터는 최소수의 템플릿 없는 FEST 웹 툴(www.stat-apps.onc.jhmi.edu/FEST) 및 "기준선 역치 무시" 파라미터로 처리되었다.

[0232] *정량화 및 통계학적 분석*

[0233] 도면의 범례에서 명확하게 언급되지 않는 한, 두 조건을 비교하는 모든 통계학적 시험은 만-휘트니 U 검정(Mann-Whitney U test)으로 수행되었다. 모든 상관관계는 피어슨 상관관계 계수로 평가되었다. 달리 언급되지 않는 한, 박스 플롯 내의 모든 박스는 분포의 중앙, 25번째 및 75번째 백분위수를 나타내고, 위스커는 10번째 및 90번째 백분위수로 연장된다. 달리 언급되지 않는 한, 모든 막대 플롯은 표준 편차(SD)를 가진 평균을 나타낸다. 플롯 및 통계학적 시험은 주로 GraphPad Prism v7.00으로 수행되었다. 모든 통계학적 시험에 대해서, ****는 $p<0.0001$을 지칭하고, ***는 $p<0.001$을 지칭하고, **는 $p<0.01$을 지칭하고, *는 $p<0.05$를 지칭한다.

[0234] **실시예 2: 정제된 조혈 선조체는 AML에서 TSA 발견을 위한 가치 있는 대조군이다.**

[0235] MS는 MAP를 직접 식별할 수 있는 유일한 사용 가능한 기술이다(Ehx and Perreault, 2019; Shao et al., 2018). 전형적으로, MAP의 MS-기반 식별은 획득한 탠덤 MS 스펙트럼을 사용자가 제공한 단백질 서열의 데이터베이스와 정합시키는 소프트웨어 툴의 사용을 통하여 수행된다. 그러나, 참조 단백질 데이터베이스는 정규 단백질 서열만 포함하고, 따라서 돌연변이 및 비정상적으로 발현된 비-정규 게놈 영역(aeTSA의 메인 소스임)으로부터 유래된 MAP의 식별을 허용하지 않는다(Laumont et al., 2018). 글로벌 TSA 식별을 위하여 맞춤화된 MS 데이터베이스를 구축하기 위한 단백질 게놈 전략은 이전에 설명된 바 있다. 맞춤형 데이터베이스는 각 종양 샘플에 대해 구축되며 다음 두 가지 기준을 충족해야 한다: 모든 잠재적인 TSA를 포함할 수 있을 만큼 충분히 포괄적이어야

하지만, 참조 데이터베이스가 부풀려지면 잘못된 발견의 위험이 증가하기 때문에 크기가 제한된다(Nesvizhskii et al., 2014; Chong et al., 2020). 데이터베이스 구축은 (i) 데이터의 핵심인 종양 샘플의 RNA-시퀀싱, (ii) RNA-seq 리드의 33개의 뉴클레오티드-길이 하위 시퀀스(k-mer)로의 인 실리코 슬라이싱, 및 (iii) 암-특이적 k-mer만을 함유하는 모듈을 작성하기 위하여 정상 k-mer의 차감으로 시작된다. 암 연구의 많은 양상과 마찬가지로, 어려운 질문은 음성 대조군(여기서는 정상 k-mer의 소스)을 선택하는 것이다. 이전의 연구에서, mTEC로부터의 k-mer는 정상 대조군으로서 사용되었다. 그러나, AML의 경우에, 다른 유형의 음성 대조군, 즉, 분류된 골수성 전구체 세포(MPC, 과립구/단핵구 선구체 및 다양한 유형의 과립구 전구체를 포함함)가 시험되었다.

[0236] mTEC 및 MPC의 값을 음성 대조군으로서 비교하기 위하여, 높은 적용 범위의 RNA-seq가 이전에 수행된(Maiga et al., 2016), 19개의 표적 AML 시편(특성에 대해서는, **표 1** 참조), 6개 mTEC 샘플 및 6개 MPC 샘플 간의 유사도가 먼저 비교되었다. 특히, MPC는 mTEC보다 AML로부터 평균 16.4% 더 많은 k-mer를 고갈시켰는데, 이는 mTEC와 AML 사이보다 MPC와 AML 사이에서 더 큰 전사체 중첩을 입증한다(**도 1A**). 따라서, mTEC와 MPC 둘 다(대략 8.7대 대략 9.9×10^{8})에 대해서 유사한 k-mer 계수치가 얻어지지만, MPC는 mTEC(대략 1.9×10^{8}, 대략 22%)보다 AML(대략 3.3×10^{8}, 대략 33%)와 더 독점적인 k-mer를 공유한다(**도 1B**). 계통의 차이가 이러한 더 높은 유사성의 기원에 있다는 것을 확립하기 위하여, 발현된 단백질-코딩 유전자의 동일성을 기반으로, 다양한 소스(방법 참조)로부터 다운로드된 분류된 상피 및 조혈 세포 RNA-seq의 어레이와 함께 AML 샘플의 t-SNE 클러스터링을 수행하였다. 이것은 AML 샘플이 조혈 세포와 함께 클러스터링된 반면 mTEC는 상피 세포와 클러스터링된 것을 나타내었다(**도 1C**). 중요하게는, mTEC는 생물학적 기능과 일치하여 가장 높은 다양성의 유전자를 발현했다(**도 1D**). 종합하면, 이들 결과는, mTEC 전사체 다양성의 다양성에도 불구하고, MPC가 AML에서 TSA의 발견에 대해 mTEC보다 더 나은 정상 대조군임을 나타낸다. 결과적으로, AML-특이적 k-mer 데이터베이스의 크기는 MPC k-mer가 mTEC k-mer 대신에 차감된 경우 더 작다.

[0237] **실시예 3: MPC-기반 TSA 발견 접근법의 개발**

[0238] 전체 AML TSA 랜드스케이프(landscape)를 캡처하는 것에 부가하여, 4가지 전략이 참조 데이터베이스 구축을 위하여 평가되었다. 처음 2가지 전략은 이전에 보고되었고(**도 2A, B**), 나머지 2가지는 새로운 것이다(**도 2C, D**). 중요하게는, AML 시편의 MS 분석은 한번만 수행되었고, 따라서 4가지 상이한 TSA-발견 접근법의 각각이 각 AML 샘플에 대해서 동일한 MS 스펙트럼에 대해 수행되었다. 첫 번째 전략은 mTEC 차감에 달려 있다(**도 2A**)(Laumont *et al.*, 2018). 두 번째는 TSA의 풍부한 소스가 될 수 있는 ERE에 의해 코딩된 MAP에 특히 중점을 둔다(**도 2B**)(Larouche *et al.*, 2020).

[0239] 세 번째 전략에서, mTEC 및 MPC 둘 다로부터의 k-mer는 고갈되었다(**도 2C**). 참고로, 고갈 단계는 k-mer의 최종 수를 콘티그 조립체에 대해서 대략 3천만으로 제한하기 위하여(더 많은 k-mer의 조립은 계산 시간의 관점에서 너무 부담이 됨) 그들의 발생(k-mer가 동일 샘플에 존재하는 횟수, **도 8a, b**)에 기반한 k-mer의 필터링에 의해 선행된다. 그 결과, mTEC+MPC k-mer의 고갈은 mTEC 단독보다 AML 샘플로부터 더 많은 k-mer를 제거하여, 발생 역치를 대략 2 내지 3배 감소시킴으로써, mTEC k-mer 고갈 접근법 데이터베이스에서 놓친 MAP의 발견을 가능하게 했다(**도 8C, D**).

[0240] 네 번째 전략은 k-mer 고갈 전략의 주요 경고인 정상 샘플 및 암 샘플에서의 k-mer 풍부도 간의 비교의 부재를 회피하는 것을 목표로 하였다. 구체적으로, k-mer 고갈 전략에서, 정상적인 대조군에서의 하나의 발생에도 불구하고, 주어진 k-mer의 존재는, 그의 빈도가 정상 대조군에 비해서 암에서 100배 더 높더라도, 암 샘플에서 이 k-mer의 필터링을 초래한다. 간단히 말해서, k-mer 발현량 차이(DKE) 분석은 DE-kupl 계산 프로토콜(Audoux *et al.*, 2017)을 사용하여 수행되고, 몇몇 인하우스 조절에 의하면, 다음과 같이 요약될 수 있다(**도 2D** 및 **도 9A**): (i) AML 샘플의 적어도 30%에서 k-mer의 사전 필터링이 존재함(발생≥3); (ii) k-mer 풍부도의 정규화; (iii) 사용자-정의된 알고리즘을 통한 k-mer 풍부도의 통계학적 비교; (iv) 유의하게 차등적으로-과발현된 k-mer(10의 최소 배수 변화)의 콘티그로의 조립 및 (v) 기원 영역을 확립시키기 위하여 게놈 상에 콘티그의 정렬. 이들은 가장 밀접하게 관련된 정상 샘플이기 때문에, MPC를 선택하여 본 연구에서 정상 대조군으로 사용했으며, 19개의 AML 샘플을 11개의 사용 가능한 고-적용범위 MPC 샘플과 비교하였다. 다음에, 개인화된 콘티그 서열을 (차등적으로 발현된 콘티그의 게놈 위치에서 리드 적용범위 및 SNP 호출에 기초하여) 각각의 AML 샘플에 대해 생성하고, 모든 가능한 리딩 프레임으로 번역하고, MAP 식별을 수행하기 위해 개인화된 정규 프로테옴과 조합하였다(**도 2**).

[0241] **실시예 4: MPC-기반 접근법은 AML에서 대부분의 TSA^{hi}를 식별한다.**

[0242] 4가지 TSA-발견 접근법의 각각은 19개 AML 샘플에 대해서 수천개의 MAP를 식별하였다(**표 2**). 실행 가능한 TSA로 간주되기 위해서는, MAP가 AML 세포에 의해 풍부하게 제시되어야 하고, 정상 세포에 의해 제시되지 않거나 T-세포 인식을 촉발하지 않을 정도로 충분히 낮은 수준으로 제시되어야 하는데, 그 이유는 에피토프 밀도가 CD8 T 세포에 의해 표적 세포의 근절에 중요한 역할을 하기 때문이다(Cosma and Eisenlohr, 2019). MAP는 우선적으로 매우 풍부한 전사체로부터 유래되기 때문에(**도 3A** 및 문헌[피어슨 et al., 2016]), 두개의 중요한 역치가 확립되었다: (i) 그 미만에서 정상 조직에서 MAP를 생성할 가능성이 낮은 것으로 간주될 수 있는 RNA 발현 수준 및 (ii) MAP를 제시할 가능성을 극적으로 증가시킬 필요가 있는 RNA 발현 배수 변화(FC). 이를 달성하기 위해, 각각의 AML 샘플에서 확인된 모든 MAP의 RNA 발현을 평가하였고, 누적 빈도 분포로 플롯팅된 정규 분포를 따르는 것으로 밝혀졌다(**도 3B**). 이것은, 8.55 리드/백만(RPHM) 미만의 발현에서, MAP을 생성할 가능성이 5%였음을 입증했다. AML 세포가 정상 과립구와 비교해서 유사한 수준의 MHC 분자를 발현하고 과립구가 정상 조직 중에서 가장 높은 수준의 MHC-I를 발현하는 점을 감안하면(Berlin *et al.*, 2015; Boegel *et al.*, 2018), 8.55 RPHM이 모든 조직에 대해서 제1 역치로서 확립되었다. 동일한 분포에 기초하여, MAP를 생성할 가능성에 대한 상이한 FC의 영향이 또한 평가될 수 있었다(**도 3C**). 이것은 2에서 5까지의 FC가 더 큰 FC보다 가능성에 더 높은 영향을 미치는 경향이 있음을 나타내었다. 따라서, 다섯(5)이 최소 FC 역치로서 채택되었다.

[0243] 이들 두 역치에 기반하여, AML, MPC, 기타 정상 조혈 세포 및 mTEC를 비롯한, 광범위한 정상 성인 조직에서 RNA 발현에 기반하여 MAP를 분리하기 위하여 결정 트리를 확립하였다(**도 3D**). 요약하면, 정상 조직에서 8.55 RPHM 미만으로 발현되고 MPC에서보다 AML에서 더 높은 수준에서 발현되는 모든 MAP는 TSA로 플래깅되는데, 그 이유는 이들의 검출이 AML 세포 표면에서 그들의 제시의 증거인 한편, 이들이 정상 조직에서 제시되는 낮은 가능성을 갖기 때문이다. 또한, AML과 MPC 사이에서 적어도 5의 FC를 갖는 TSA는 TSA^{hi}로 플래깅되는, 그 이유는 이들이 AML 세포에 의해 배타적으로 제시되는 최고 가능성을 갖기 때문이다. 다른 조직에 비해서 조혈 세포에서 과발현되지만 이러한 기준을 충족시키는데 실패한 다른 MAP는 TAA 또는 조혈 특이적 항원(HSA)으로 분류되었다(**도 3D**).

[0244] 각 파이프라인의 사전 필터링 단계(방법 참조) 및 결정 트리에 기반한 범주화 후에, 관심 MAP(MOI)의 4가지 목록이 얻어졌다(**표 2**). mTEC 고갈 접근법은 가장 높은 비율의 HSA를 산출한 반면 대다수의 TSA^{hi}가 MPC-기반 접근법에 의해 식별되었다(**도 3E, F** 및 **도 10A**). DKE 접근법은 최소한의 환자에서 최소 발생을 갖는 k-mer를 사전 필터링하기 때문에, 두 MPC-기반 접근법 간의 중첩이 낮았다. 따라서, 고갈 접근법에 의해 식별된 대부분의 TSA^{hi}는 DKE 접근법에 의해 식별된 것과 비교하여 낮은 환자간 공유를 제시하였다(**도 9B**). 종합하면, 이들 결과는 DKE 접근법이 AML에서 TSA^{hi}를 식별하는데 가장 적합하고, 추가로 덜 공유된 TSA^{hi}를 식별하기 위해 MPC-기반 k-mer 고갈 접근법에 의해 보완될 수 있는 것을 나타낸다.

[0245] [표 2]

4 가지 프로테오게노믹 접근법을 통해서 식별된 MOI 의 상세.

MOI 서열(서열번호)	분류	식별 방법	생물형	유전자(엑손 또는 인트론인지의 여부)
RQISVQASL (1)	HSA	K-mer 발현량 차이	인트론	PPP6R2
DRELRNLEL (2)	HSA	K-mer 발현량 차이	코딩 엑손	PYHIN1
GARQQIHSW (3)	HSA	K-mer 발현량 차이	코딩 엑손	FADS1
SGKLRVAL (4)	HSA	K-mer 발현량 차이	코딩 엑손	TBX21
RSASSATQVHK (5)	HSA	K-mer 발현량 차이	유전자간	
SASSATQVHK (6)	HSA	K-mer 발현량 차이	유전자간	
FLLEFKPVS (7)	HSA	K-mer 발현량 차이	인트론	CEP83
GPQVRGSI (8)	HSA	K-mer 발현량 차이	인트론	UBE3C
IRMKAQAL (9)	HSA	K-mer 발현량 차이	인트론	KMT2C
KIKVFSKVY (10)	HSA	K-mer 발현량 차이	인트론	TRPM7
LLSRGLLFRI (11)	HSA	K-mer 발현량 차이	인트론	ANKIB1
LPIASASLL (12)	HSA	K-mer 발현량 차이	인트론	MAML3
LYFLGHGSI (13)	HSA	K-mer 발현량 차이	인트론	KMT2C
NPLQLSLSI (14)	HSA	K-mer 발현량 차이	인트론	ARL15
DLMLRESL (15)	HSA	K-mer 발현량 차이	ncRNA	RP11-477N3.1
VTFKLSLF (16)	HSA	K-mer 발현량 차이 및 ERE	ERE	
IALYKQVL (17)	HSA	ERE	ERE	
IVATGSLLK (18)	HSA	ERE	ERE	
KIKNKTKNK (19)	HSA	ERE	ERE	
KLLSLTIYK (20)	HSA	ERE	ERE	
NILKKTVL (21)	HSA	ERE	ERE	
NPKLKDIL (22)	HSA	ERE	ERE	
NQKKVRIL (23)	HSA	ERE	ERE	
PFPLVQVEPV (24)	HSA	ERE	ERE	
SPQSGPAL (25)	HSA	ERE	ERE	
TSRLPKIQK (26)	HSA	ERE	ERE	
LLDNILQSI (27)	HSA	ERE 및 mTEC k-mer 고갈	ERE	
RLEVRKVIL (28)	HSA	mTEC k-mer 고갈	엑손-인트론	DOCK8
LSWGYFLFK (29)	HSA	mTEC k-mer 고갈	유전자간	
TILPRILTL (30)	HSA	mTEC k-mer 고갈	유전자간	
EGKIKRNI (31)	HSA	mTEC k-mer 고갈	인트론	SLC25A13
FLASFVEKTVL (32)	HSA	mTEC k-mer 고갈	인트론	OS9
ILASHNLTV (33)	HSA	mTEC k-mer 고갈	인트론	RPS6KC1

[0246]

IQLTSVHLL (34)	HSA	mTEC k-mer 고갈	인트론	PTPRC
LELISFLPVL (35)	HSA	mTEC k-mer 고갈	인트론	CDCA8
NFCMLHQSI (36)	HSA	mTEC k-mer 고갈	인트론	IKZF1
PARPAGPL (37)	HSA	mTEC k-mer 고갈	인트론	INPP5D
PLPIVPAL (38)	HSA	mTEC k-mer 고갈	인트론	USP3
SNLIRTGSH (39)	HSA	mTEC k-mer 고갈	인트론	INPP5D
VPAPAQAI (40)	HSA	mTEC k-mer 고갈	인트론	PAN3
KGHGGPRSW (41)	HSA	mTEC k-mer 고갈	ncRNA	AC104232.1
ITSSAVTTALK (42)	HSA	mTEC k-mer 고갈	UTR	LMO2
LLLPESPSI (43)	HSA	mTEC+MPC k-mer 고갈 및 ERE 및 mTEC k-mer 고갈	ERE	
VILIPLPPK (44)	HSA	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	ERE 안티센스	
AVLLPKPPK (45)	HSA	mTEC+MPC k-mer 고갈	ERE 안티센스	
TQVSMAESI (46)	HSA	mTEC+MPC k-mer 고갈	ERE	
LNHLRTSI (47)	HSA	mTEC+MPC k-mer 고갈	인트론	ZNF407
NTSHLPLIY (48)	TAA	K-mer 발현량 차이	인트론	COL4A5
SIQRNLSL (49)	TAA	K-mer 발현량 차이	코딩 엑손	PIEZO2
NVSSHVHTV (50)	TAA	K-mer 발현량 차이	코딩 엑손	TPSB2
ALASHLIEA (51)	TAA	K-mer 발현량 차이	코딩 엑손	EHD2
ALDDITIQL (52)	TAA	K-mer 발현량 차이	코딩 엑손	MAMDC2
ALGNTVPAV (53)	TAA	K-mer 발현량 차이	코딩 엑손	DNAAF3
ALLPAVPSL (54)	TAA	K-mer 발현량 차이	코딩 엑손	WT1
APAPPPVAV (55)	TAA	K-mer 발현량 차이	코딩 엑손	IRX3
APDKKITL (56)	TAA	K-mer 발현량 차이	코딩 엑손	FOXC1
AQMNLLQKY (57)	TAA	K-mer 발현량 차이	코딩 엑손	DLC1
DQVIRLAGL (58)	TAA	K-mer 발현량 차이	코딩 엑손	PTPN14
ETTSQVRKY (59)	TAA	K-mer 발현량 차이	코딩 엑손	MMRN1
GGSLIHPQW (60)	TAA	K-mer 발현량 차이	코딩 엑손	TPSB2
GLYYKLHNV (61)	TAA	K-mer 발현량 차이	코딩 엑손	GATA2
GQKPVILTY (62)	TAA	K-mer 발현량 차이	코딩 엑손	IGSF10
GSLDFQRGW (63)	TAA	K-mer 발현량 차이	코딩 엑손	ANGPT1
HHLVETLKF (64)	TAA	K-mer 발현량 차이	코딩 엑손	GGT5
HLLSETPQL (65)	TAA	K-mer 발현량 차이	코딩 엑손	IGSF10
HQLYRASAL (66)	TAA	K-mer 발현량 차이	코딩 엑손	TTC28
HTDDIENAKY (67)	TAA	K-mer 발현량 차이	코딩 엑손	PTPN14
IAAPILHV (68)	TAA	K-mer 발현량 차이	코딩 엑손	GGT5
KAFPFHIIF (69)	TAA	K-mer 발현량 차이	코딩 엑손	GUCY1B3
KATEYVHSL (70)	TAA	K-mer 발현량 차이	코딩 엑손	MYCN
KFSNVTMLF (71)	TAA	K-mer 발현량 차이	코딩 엑손	GUCY1A3
KLLEKAFSI (72)	TAA	K-mer 발현량 차이	코딩 엑손	CYP7B1
KPMPTKVVF (73)	TAA	K-mer 발현량 차이	코딩 엑손	MAMDC2
NVNRPLTMK (74)	TAA	K-mer 발현량 차이	코딩 엑손	GATA2
REPYELTVPAL (75)	TAA	K-mer 발현량 차이	코딩 엑손	RBPMS
SEAEAAKNAL (76)	TAA	K-mer 발현량 차이	코딩 엑손	RBPMS
SLWGQPAEA (77)	TAA	K-mer 발현량 차이	코딩 엑손	COL4A5
SPADHRGYASL (78)	TAA	K-mer 발현량 차이	코딩 엑손	SOX4
SPQSAAAEL (79)	TAA	K-mer 발현량 차이	코딩 엑손	FOXC1
SPVVHQSL (80)	TAA	K-mer 발현량 차이	코딩 엑손	PTPN14
SPYRTPVL (81)	TAA	K-mer 발현량 차이	코딩 엑손	IGSF10
SVFAGVVGV (82)	TAA	K-mer 발현량 차이	코딩 엑손	GUCY1A3
SYSPAHARL (83)	TAA	K-mer 발현량 차이	코딩 엑손	GATA2
THGSEQLHL (84)	TAA	K-mer 발현량 차이	코딩 엑손	IGSF10
TQAPPNVVL (85)	TAA	K-mer 발현량 차이	코딩 엑손	DENND6B
VLVPYEPPQV (86)	TAA	K-mer 발현량 차이	코딩 엑손	TP63
VSFPDVRKV (87)	TAA	K-mer 발현량 차이	코딩 엑손	NEGR1
VVFDKSDLAKY (88)	TAA	K-mer 발현량 차이	코딩 엑손	SLC45A3

[0247]

YSHHSGLEY (89)	TAA	K-mer 발현량 차이	코딩 엑손	CAV2
YYLDWIHHY (90)	TAA	K-mer 발현량 차이	코딩 엑손	TPSB2
SVYKYLKAK (91)	TAA	K-mer 발현량 차이	UTR	IRX3
IYQFIMDRF (92)	TAA	K-mer 발현량 차이	코딩 엑손	FOXC1
GTLQGIRAW (93)	TAA	K-mer 발현량 차이 및 mTEC+MPC k-mer 고갈	코딩 엑손	NKX2-3
AQKVSVGQAA (94)	TAA	mTEC k-mer 고갈	UTR	LMO2
LYPSKLTHF (95)	TAA	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	유전자간	
ATQNTIIGK (96)	TAA	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	프레임 밖 번역	CPA3
AQDIILQAV (97)	TSA^{hi}	K-mer 발현량 차이	유전자간	
PPRPLGAQV (98)	TSA^{hi}	K-mer 발현량 차이	UTR	C16orf87
FNVALNARY (99)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	LTBP1
GPGSRESTL (100)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	PLPPR3
IPHQRSSL (101)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	GCSAML
LTDRIYLTL (102)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	DNAH10
NLKEKKALF (103)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	ST8SIA6
VLFGGKVSGA (104)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	MFSD2B
VVFPFPVNK (105)	TSA^{hi}	K-mer 발현량 차이	코딩 엑손	MYCN
SLLIIPKKK (106)	TSA^{hi}	K-mer 발현량 차이	ERE	
APGAAGQRL (107)	TSA^{hi}	K-mer 발현량 차이	유전자간	
KLQDKEIGL (108)	TSA^{hi}	K-mer 발현량 차이	유전자간	
SLREPQPAL (109)	TSA^{hi}	K-mer 발현량 차이	유전자간	
TPGRSTQAI (110)	TSA^{hi}	K-mer 발현량 차이	유전자간	
APRGTAAL (111)	TSA^{hi}	K-mer 발현량 차이	인트론	JADE1
IASPIALL (112)	TSA^{hi}	K-mer 발현량 차이	인트론	ELL
ILFQNSALK (113)	TSA^{hi}	K-mer 발현량 차이	인트론	SLC25A25
ILKKNISI (114)	TSA^{hi}	K-mer 발현량 차이	인트론	AKAP13
IPLAVRTI (115)	TSA^{hi}	K-mer 발현량 차이	인트론	FBXO28
LPRNKPLL (116)	TSA^{hi}	K-mer 발현량 차이	인트론	IGF2BP2
PAPPHPAAL (117)	TSA^{hi}	K-mer 발현량 차이	인트론	HOOK2
SPVVRVGL (118)	TSA^{hi}	K-mer 발현량 차이	인트론	DNAH7
TLNQGINVYI (119)	TSA^{hi}	K-mer 발현량 차이	인트론	SLC39A10
RPRGPRTAP (120)	TSA^{hi}	K-mer 발현량 차이	UTR	PIEZO1
SVQLLEQAIHK (121)	TSA^{hi}	K-mer 발현량 차이 및 ERE	ERE	
RTPKNYQHW (122)	TSA^{hi}	K-mer 발현량 차이 및 mTEC k-mer 고갈	코딩 엑손	LINC01835
ALPVALPSL (123)	TSA^{hi}	K-mer 발현량 차이 및 mTEC k-mer 고갈	인트론	RFX2
SLQILVSSL (124)	TSA^{hi}	K-mer 발현량 차이 및 mTEC k-mer 고갈	인트론	LRRC8C
ISNKVPKLF (125)	TSA^{hi}	K-mer 발현량 차이 및 mTEC k-mer 고갈	ncRNA	LINC02147
TVIRIAIVNK (126)	TSA^{hi}	K-mer 발현량 차이 및 mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	인트론	FTO
KEIFLELRL (127)	TSA^{hi}	ERE	ERE	
TLRSPGSSL (128)	TSA^{hi}	ERE	ERE	
TVRGDVSSL (129)	TSA^{hi}	mTEC k-mer 고갈	유전자간	
ALDPLLLRI (130)	TSA^{hi}	mTEC k-mer 고갈	인트론	CETP
ISLIVTGLK (131)	TSA^{hi}	mTEC k-mer 고갈	인트론	ASPH
KILDVNLRI (132)	TSA^{hi}	mTEC k-mer 고갈	인트론	SIK3
ERVYIRASL (133)	TSA^{hi}	mTEC k-mer 고갈	ncRNA	AL359636.1
ILDLESRY (134)	TSA^{hi}	mTEC+MPC k-mer 고갈	ERE	
KTFVQQKTL (135)	TSA^{hi}	mTEC+MPC k-mer 고갈	ERE	
LYIKSLPAL (136)	TSA^{hi}	mTEC+MPC k-mer 고갈	ERE	

[0248]

[0249]

VLKEKNASL (137)	TSA^{hi}	mTEC+MPC k-mer 고갈	ERE	
LGISLTLKY (138)	TSA^{hi}	mTEC+MPC k-mer 고갈	유전자간	
DLLPKKLL (139)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	BACH1
HSLISIVYL (140)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	XACT
IAGALRSVL (141)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	TNS3
IGNPILRVL (142)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	EFR3A
IYAPHIRLS (143)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	ERC1
LRSQILSY (144)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	LINC00484
RYLANKIHI (145)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	ARHGAP32
SLLSGLLRA (146)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	CD34
SRIHLVVL (147)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	ZNF280C
SSSPVRGPSV (148)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	PTK2
STFSLYLKK (149)	TSA^{hi}	mTEC+MPC k-mer 고갈	인트론	DTNA
SLDLLPLSI (150)	TSA^{hi}	mTEC+MPC k-mer 고갈	ncRNA	LINC00996
VTDLLALTV (151)	TSA^{hi}	mTEC+MPC k-mer 고갈 및 ERE 및 mTEC k-mer 고갈	ERE	
RTQITKVSLKK (152)	TSA^{hi}	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	유전자간	
ILRSPLKW (153)	TSA^{hi}	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	인트론	PIEZO2
LSTGHLSTV (154)	TSA^{hi}	mTEC+MPC k-mer 고갈 및 mTEC k-mer 고갈	인트론	MYO16
TVEEYLVNI (155)	TSA^{lo}	K-mer 발현량 차이	인트론	ECSIT
QIKTKLLGSL (156)	TSA^{lo}	K-mer 발현량 차이	인트론	KIF27
LPSFSHFLLL (157)	TSA^{lo}	K-mer 발현량 차이	유전자간	
CLRIGPVTL (158)	TSA^{lo}	K-mer 발현량 차이	인트론	ZDHHC14
HVSDGSTALK (159)	TSA^{lo}	K-mer 발현량 차이	인트론	PDE4D
IAYSVRALR (160)	TSA^{lo}	K-mer 발현량 차이	인트론	ACBD3
PRGFLSAL (161)	TSA^{lo}	K-mer 발현량 차이	인트론	ZNF26
ISSWLISSL (162)	TSA^{lo}	K-mer 발현량 차이 및 ERE	ERE	
IPLNPFSSL (163)	TSA^{lo}	K-mer 발현량 차이 및 mTEC k-mer 고갈	ERE	
LSDRQLSL (164)	TSA^{lo}	K-mer 발현량 차이 및 mTEC k-mer 고갈	인트론	CDK6
LSHPAPSSL (165)	TSA^{lo}	K-mer 발현량 차이 및 mTEC k-mer 고갈	인트론	RNF220
LRKAVDPIL (166)	TSA^{lo}	K-mer 발현량 차이 및 mTEC+MPC k-mer 고갈	인트론	SLC39A11
ILLEEQSLI (167)	TSA^{lo}	ERE	ERE	
LTSISIRPV (168)	TSA^{lo}	ERE	ERE	
TISECPLLI (169)	TSA^{lo}	ERE	ERE	
TLKLKKIFF (170)	TSA^{lo}	ERE	ERE	
ILLSNFSSL (171)	TSA^{lo}	mTEC k-mer 고갈	ERE	
LGGAWKAVF (172)	TSA^{lo}	mTEC k-mer 고갈	ERE	
LSASHLSSL (173)	TSA^{lo}	mTEC k-mer 고갈	ERE	
AGDIIARLI (174)	TSA^{lo}	mTEC k-mer 고갈	인트론	ZWILCH
DRGILRNLL (175)	TSA^{lo}	mTEC k-mer 고갈	인트론	ABI1
GLRLIHVSL (176)	TSA^{lo}	mTEC k-mer 고갈	인트론	KLF13
GLRLLHVSL (177)	TSA^{lo}	mTEC k-mer 고갈	인트론	VAV3
LHNEKGLSL (178)	TSA^{lo}	mTEC k-mer 고갈	인트론	ZNF804A
LPSFSRPSGII (179)	TSA^{lo}	mTEC k-mer 고갈	인트론	CCDC26
LSSRLPLGK (180)	TSA^{lo}	mTEC k-mer 고갈	인트론	ZNF367
MIGIKRLL (181)	TSA^{lo}	mTEC k-mer 고갈	인트론	TRIP12
NLKKREIL (182)	TSA^{lo}	mTEC k-mer 고갈	인트론	POGLUT1
RMVAYLQQL (183)	TSA^{lo}	mTEC k-mer 고갈	인트론	PTEN
SPARALPSL (184)	TSA^{lo}	mTEC k-mer 고갈	인트론	RFX8

[0250]

TVPGIQRY (185)	TSA^{lo}	mTEC k-mer 고갈	인트론	ERBIN
VSRNYVLLI (186)	TSA^{lo}	mTEC k-mer 고갈	인트론	SLC39A10
LTVPLSVFW (187)	TSA^{lo}	mTEC k-mer 고갈	위유전자 (Pseudogene)	SMURF2P1-LRRC37BP1
KLNQAFLVL (188)	TSA^{lo}	mTEC k-mer 고갈	인트론	SNTB1
RLVSSTLLQK (189)	TSA^{lo}	mTEC+MPC k-mer 고갈 및 ERE 및 mTEC k-mer 고갈	ncRNA	AC128707.1
LPSHSLLI (190)	TSA^{lo}	mTEC+MPC k-mer 고갈	인트론	MAPRE2

[0251]

MOI	면역원성 스코어	AML/MPC	샘플에서 식별됨	HLA
RQISVQASL (1)	0,3896	0,143096844	15H080	HLA-B*27:05
DRELRNLEL (2)	0,36072	109,6504772	08H039	HLA-B*14:02
GARQQIHSW (3)	0,31568	8,852871755	07H122	HLA-B*57:01
SGKLRVAL (4)	0,38024	277,184	05H143, 07H060	HLA-B*08:01
RSASSATQVHK (5)	0,27736	21,15087756	15H023	HLA-A*03:01
SASSATQVHK (6)	0,33728	22,04593113	08H053	HLA-A*11:01
FLLEFKPVS (7)	0,32488	0,178033826	14H124	HLA-A*02:01
GPQVRGSI (8)	0,43896	0,244436854	07H063	HLA-B*07:02
IRMKAQAL (9)	0,30384	0,14128544	07H063	HLA-C*06:02
KIKVFSKVY (10)	0,15168	0,563808415	12H172	HLA-B*15:01
LLSRGLLFRI (11)	0,40744	0,337716083	14H124	HLA-A*02:01
LPIASASLL (12)	0,50848	0,941520982	15H023	HLA-B*51:01
LYFLGHGSI (13)	0,3268	0,254931815	16H123	HLA-A*24:02
NPLQLSLSI (14)	0,45888	0,488543537	07H060	HLA-B*08:01
DLMLRESL (15)	0,4704	0,543741144	07H060	HLA-B*08:01
VTFKLSLF (16)	0,38384	0,890019017	05H143	HLA-B*57:01
IALYKQVL (17)	0,37872	0,841619533	05H143	HLA-B*08:01
IVATGSLLK (18)	0,406	0,506704612	15H023	HLA-A*03:01
KIKNKTKNK (19)	0,386	0,846822574	16H123	HLA-A*03:01
KLLSLTIYK (20)	0,2392	0,360396852	16H123	HLA-A*03:01
NILKKTVL (21)	0,42712	0,995471668	15H063	HLA-B*08:01
NPKLKDIL (22)	0,37752	1,97918996	05H143	HLA-B*08:01
NQKKVRIL (23)	0,33296	0,456777709	07H060	HLA-B*08:01
PFPLVQVEPV (24)	0,48656	0,406411352	15H023	HLA-B*51:01
SPQSGPAL (25)	0,44536	0,540664564	16H123	HLA-B*07:02
TSRLPKIQK (26)	0,29104	0,536699874	07H063	HLA-A*30:01
LLDNILQSI (27)	0,42688	2,344739646	15H023	HLA-A*02:01
RLEVRKVIL (28)	0,2716	0,619722353	15H063	HLA-B*08:01
LSWGYFLFK (29)	0,27096	2,09747864	07H063	HLA-A*30:01
TILPRILTL (30)	0,4464	0,603930045	16H123	HLA-C*07:02
EGKIKRNI (31)	0,36296	0,79453462	15H063	HLA-B*08:01
FLASFVEKTVL (32)	0,2804	0,317704238	14H124	HLA-A*02:01
ILASHNLTV (33)	0,35944	0,836592866	16H145	HLA-A*02:01
IQLTSVHLL (34)	0,37928	0,364630795	15H080	HLA-A*02:01
LELISFLPVL (35)	0,42824	0,089927659	14H124	HLA-A*02:01
NFCMLHQSI (36)	0,42648	0,548351918	12H172	HLA-A*24:02
PARPAGPL (37)	0,44312	0,830983729	16H145	HLA-C*03:03
PLPIVPAL (38)	0,51424	0,60208428	15H023	HLA-B*51:01
SNLIRTGSH (39)	0,29392	0,834424495	08H039	HLA-B*14:02
VPAPAQAI (40)	0,462	0,085637644	16H123	HLA-B*07:02
KGHGGPRSW (41)	0,25872	0,352735092	07H122	HLA-B*57:01
ITSSAVTTALK (42)	0,37744	2,017525959	11H008, 15H023, 16H123	HLA-A*03:01
LLLPESPSI (43)	0,39184	0,178058578	08H039	HLA-A*02:01
VILIPLPPK (44)	0,40664	0,262280081	11H035	HLA-A*03:01

AVLLPKPPK (45)	0,35024	0,545425585	10H005	HLA-A*11:01
TQVSMAESI (46)	0,34456	0,759898781	07H141	HLA-B*38:01
LNHLRTSI (47)	0,4288	0,275551012	05H143	HLA-B*08:01
NTSHLPLIY (48)	0,24416	405,694	07H060	HLA-A*01:01
SIQRNLSL (49)	0,32568	19,18764464	07H060	HLA-B*08:01
NVSSHVHTV (50)	0,27864	7,343047581	05H149	HLA-A*68:02
ALASHLIEA (51)	0,07824	15,39591183	08H039, 15H063	HLA-A*02:01
ALDDITIQL (52)	0,36416	17,31517663	05H143	HLA-A*02:01
ALGNTVPAV (53)	0,4856	14,04712651	07H122	HLA-A*02:01
ALLPAVPSL (54)	0,47288	1261,102	05H143, 07H122, 16H145	HLA-A*02:01
APAPPPVAV (55)	0,53488	344,1774537	16H123	HLA-B*07:02
APDKKITL (56)	0,2792	55,02751683	11H035	HLA-B*07:02
AQMNLLQKY (57)	0,26056	31,38609992	12H172	HLA-B*15:01
DQVIRLAGL (58)	0,44592	246,7135986	08H039	HLA-B*14:02
ETTSQVRKY (59)	0,15824	9,473637252	07H141, 14H124	HLA-A*26:01
GGSLIHPQW (60)	0,45512	11,49478271	15H013	HLA-B*57:03
GLYYKLHNV (61)	0,23424	13,52738214	07H122	HLA-A*02:01
GQKPVILTY (62)	0,21208	5,212085157	16H123	HLA-B*15:01
GSLDFQRGW (63)	0,31896	19,21014006	07H122	HLA-B*57:01
HHLVETLKF (64)	0,23928	66,82447284	07H141	HLA-B*38:01
HLLSETPQL (65)	0,51952	13,50509609	05H143, 16H145	HLA-A*02:01
HQLYRASAL (66)	0,42744	9,735569575	08H039	HLA-B*14:02
HTDDIENAKY (67)	0,08944	88,6346121	08H039	HLA-A*01:01
IAAPILHV (68)	0,45152	26,19480016	15H023	HLA-B*51:01
KAFPFHIIF (69)	0,31792	8,766274199	07H122	HLA-B*57:01
KATEYVHSL (70)	0,14792	67,33534262	11H008	HLA-C*06:02
KFSNVTMLF (71)	0,39376	8,927105041	05H149	HLA-A*24:02
KLLEKAFSI (72)	0,28016	20,37743121	05H143, 07H122	HLA-A*02:01
KPMPTKVVF (73)	0,32808	20,74788406	11H035	HLA-B*07:02
NVNRPLTMK (74)	0,28872	12,365496	11H008, 16H123	HLA-A*03:01
REPYELTVPAL (75)	0,39344	8,664178443	08H039	HLA-B*40:01
SEAEAAKNAL (76)	0,19208	15,84619972	08H039	HLA-B*40:01
SLWGQPAEA (77)	0,89872	249,8139721	15H063, 16H145	HLA-A*02:01
SPADHRGYASL (78)	0,31104	28,5105523	11H035, 16H123	HLA-B*07:02
SPQSAAAEL (79)	0,32288	59,61837148	11H035, 16H123	HLA-B*07:02
SPVVHQSL (80)	0,4292	52,97700997	11H035	HLA-B*07:02
SPYRTPVL (81)	0,4632	8,927639423	07H063, 11H035	HLA-B*07:02
SVFAGVVGV (82)	0,112	8,69977003	05H143, 16H145	HLA-A*02:01
SYSPAHARL (83)	0,2644	8,83718913	11H035, 16H123	HLA-C*07:02
THGSEQLHL (84)	0,3024	3,293357687	07H141	HLA-B*38:01
TQAPPNVVL (85)	0,44232	11,85852997	15H013	HLA-C*07:02
VLVPYEPPQV (86)	0,48944	641,852	07H122	HLA-A*02:01
VSFPDVRKV (87)	0,2852	9,647921856	05H143	HLA-C*06:02
VVFDKSDLAKY (88)	0,16416	9,216067541	07H060	HLA-A*29:02
YSHHSGLEY (89)	0,246	91,16567732	08H053	HLA-A*01:01
YYLDWIHHY (90)	0,29696	21,43744849	15H013, 16H123	HLA-C*07:02
SVYKYLKAK (91)	0,19256	8068,72	11H008, 16H123	HLA-A*03:01
IYQFIMDRF (92)	0,25448	86,11916983	16H123	HLA-A*24:02
GTLQGIRAW (93)	0,37888	1098,148	05H143, 11H008	HLA-B*57:01
AQKVSVGQAA (94)	0,24528	2,691338811	12H172	HLA-B*15:01
LYPSKLTHF (95)	0,22144	5,627619975	12H172	HLA-A*24:02
ATQNTIIGK (96)	0,22592	2,541628072	08H053	HLA-A*11:01
AQDIILQAV (97)	0,34352	230,14	08H039	HLA-C*08:02
PPRPLGAQV (98)	0,4844	14,35433865	07H063	HLA-B*07:02
FNVALNARY (99)	0,29216	113,6001075	07H060, 11H008	HLA-A*29:02
GPGSRESTL (100)	0,41576	19,38141684	16H123	HLA-B*07:02

[0252]

[0253]

IPHQRSSL (101)	0,37912	61,72974486	15H063	HLA-B*08:01
LTDRIYLTL (102)	0,34256	203,504	08H039	HLA-C*08:02
NLKEKKALF (103)	0,33608	8,968146067	07H060	HLA-B*08:01
VLFGGKVSGA (104)	0,30776	18,89924792	07H122	HLA-A*02:01
VVFPFPVNK (105)	0,34848	1449,166	11H035, 16H123	HLA-A*03:01
SLLIIPKKK (106)	0,38096	5,215809797	10H005	HLA-A*11:01
APGAAGQRL (107)	0,42808	24,61786022	07H063, 11H035	HLA-B*07:02
KLQDKEIGL (108)	0,24288	343,936	16H145	HLA-A*02:01
SLREPQPAL (109)	0,50512	69,20553879	15H013	HLA-C*07:02
TPGRSTQAI (110)	0,39992	8,314753382	16H123	HLA-B*07:02
APRGTAAL (111)	0,45936	17,29079131	11H035	HLA-B*07:02
IASPIALL (112)	0,44496	20,11176673	12H172	HLA-C*03:03
ILFQNSALK (113)	0,33824	16,92592279	16H123	HLA-A*03:01
ILKKNISI (114)	0,37568	22,53624085	07H060	HLA-B*08:01
IPLAVRTI (115)	0,4452	8,644117906	15H023	HLA-B*51:01
LPRNKPLL (116)	0,36928	6,213173676	15H023	HLA-B*51:01
PAPPHPAAL (117)	0,43568	13,04546221	07H063	HLA-C*07:02
SPVVRVGL (118)	0,45504	22,15400444	16H123	HLA-B*07:02
TLNQGINVYI (119)	0,2868	9,544366847	14H124	HLA-A*02:01
RPRGPRTAP (120)	0,5112	29,95425988	07H063, 11H035	HLA-B*07:02
SVQLLEQAIHK (121)	0,33208	56,5275384	10H005	HLA-A*11:01
RTPKNYQHW (122)	0,2	53,71748686	07H122	HLA-B*57:01
ALPVALPSL (123)	0,55128	16,47119814	16H145, 14H124	HLA-A*02:01
SLQILVSSL (124)	0,414	6,710019262	08H039	HLA-B*14:02
ISNKVPKLF (125)	0,29072	19,42035741	05H143	HLA-B*57:01
TVIRIAIVNK (126)	0,35448	43568,4	16H123	HLA-A*03:01
KEIFLELRL (127)	0,29152	90,9946	11H008	HLA-B*44:03
TLRSPGSSL (128)	0,2984	11,27169102	11H035	HLA-B*07:02
TVRGDVSSL (129)	0,23032	9,688689585	11H035	HLA-B*07:02
ALDPLLLRI (130)	0,4948	9,563532996	07H122	HLA-A*02:01
ISLIVTGLK (131)	0,39056	9,244635345	15H023	HLA-A*03:01
KILDVNLRI (132)	0,28152	8,266412385	15H023	HLA-A*02:01
ERVYIRASL (133)	0,24856	7,705323592	08H039	HLA-B*14:02
ILDLESRY (134)	0,3588	16,434	07H063	HLA-A*01:01
KTFVQQKTL (135)	0,1732	261,178	07H122	HLA-B*57:01
LYIKSLPAL (136)	0,2556	19,06117834	08H039	HLA-B*14:02
VLKEKNASL (137)	0,2932	75,9224	05H143	HLA-B*08:01
LGISLTLKY (138)	0,29328	6,31266165	07H060	HLA-A*29:02
DLLPKKLL (139)	0,40048	141,2948	07H060	HLA-B*08:01
HSLISIVYL (140)	0,25104	80,1734	16H145	HLA-C*03:03
IAGALRSVL (141)	0,4288	221,362	08H039	HLA-B*14:02
IGNPILRVL (142)	0,5164	95,0262	05H143, 07H122	HLA-C*07:01, HLA-C*06:02
IYAPHIRLS (143)	0,32096	230,392	05H143	HLA-C*07:01
LRSQILSY (144)	0,28312	33,62998983	15H080	HLA-B*27:05
RYLANKIHI (145)	0,16208	64,5122	05H149	HLA-A*24:02
SLLSGLLRA (146)	0,5932	308,14	07H122	HLA-A*02:01
SRIHLVVL (147)	0,4008	60,2848	05H143	HLA-B*08:01
SSSPVRGPSV (148)	0,37792	113,1364	05H149	HLA-A*68:02
STFSLYLKK (149)	0,18904	98,973	10H005	HLA-A*11:01
SLDLLPLSI (150)	0,39032	17,23616137	07H141, 08H039	HLA-C*05:01, HLA-A*02:01
VTDLLALTV (151)	0,42864	10,81032029	08H039	HLA-A*01:01
RTQITKVSLKK (152)	0,21872	2051,04	08H053	HLA-A*11:01
ILRSPLKW (153)	0,41552	56,7804	07H122	HLA-B*57:01
LSTGHLSTV (154)	0,32176	45,7274	07H122	HLA-C*06:02
TVEEYLVNI (155)	0,24264	1,911770046	15H063	HLA-C*07:01

QIKTKLLGSL (156)	0,3024	1,404844942	07H060	HLA-B*08:01
LPSFSHFLLL (157)	0,38896	2,7505537	07H063	HLA-B*07:02
CLRIGPVTL (158)	0,54248	2,362675236	11H035	HLA-C*07:02
HVSDGSTALK (159)	0,20528	2,419142618	15H023	HLA-A*03:01
IAYSVRALR (160)	0,38472	3,755291737	11H035	HLA-A*03:01
PRGFLSAL (161)	0,45264	4,511422509	07H063	HLA-B*07:02
ISSWLISSL (162)	0,378	4,706302579	08H039	HLA-B*14:02
IPLNPFSSL (163)	0,3876	2,67504273	16H123	HLA-B*07:02
LSDRQLSL (164)	0,42816	1,037298278	07H060	HLA-A*01:01
LSHPAPSSL (165)	0,4132	2,060547234	07H063	HLA-A*30:01
LRKAVDPIL (166)	0,32448	3,901328114	07H063	HLA-C*06:02
ILLEEQSLI (167)	0,4304	4,359584524	16H145	HLA-A*02:01
LTSISIRPV (168)	0,39904	1,503160093	08H039	HLA-A*02:01
TISECPLLI (169)	0,52632	1,643501465	08H039	HLA-A*02:01
TLKLKKIFF (170)	0,23936	2,527260887	05H143	HLA-B*08:01
ILLSNFSSL (171)	0,33432	1,032824295	15H023	HLA-A*02:01
LGGAWKAVF (172)	0,38584	1,397521024	15H013	HLA-B*57:03
LSASHLSSL (173)	0,37704	2,203674156	07H141	HLA-C*12:03
AGDIIARLI (174)	0,40696	1,817220719	08H039	HLA-C*08:02
DRGILRNLL (175)	0,46584	1,753521623	08H039	HLA-B*14:02
GLRLIHVSL (176)	0,37616	2,738686931	08H039	HLA-B*14:02
GLRLLHVSL (177)	0,388	3,327626689	08H039	HLA-B*14:02
LHNEKGLSL (178)	0,23976	1,21912196	05H143	HLA-C*07:01
LPSFSRPSGII (179)	0,38304	2,058642372	07H063	HLA-B*07:02
LSSRLPLGK (180)	0,32344	3,016124162	11H008	HLA-A*03:01
MIGIKRLL (181)	0,4116	2,454205925	05H143	HLA-B*08:01
NLKKREIL (182)	0,42808	1,722302438	07H060	HLA-B*08:01
RMVAYLQQL (183)	0,41624	1,99560151	12H172	HLA-A*02:01
SPARALPSL (184)	0,51256	1,404210958	16H123	HLA-B*07:02
TVPGIQRY (185)	0,39184	1,304086581	14H124	HLA-A*26:01
VSRNYVLLI (186)	0,24664	4,908161328	07H060	HLA-C*07:01
LTVPLSVFW (187)	0,40472	1,193403976	05H143	HLA-B*57:01
KLNQAFLVL (188)	0,35824	1,82047525	15H080	HLA-A*02:01
RLVSSTLLQK (189)	0,32288	1,091198313	11H035	HLA-A*03:01
LPSHSLLI (190)	0,39184	2,067938041	15H023	HLA-B*51:01

생물형은 표시된 게놈 위치에서 펩타이드 코딩 서열의 조사 시에 수동적으로 귀속되었다. 면역원성 스코어는 레피토프(Repitope)로 계산되었다. HLA 대립유전자는, netMHC4.0에 의해 예측되는 바와 같이, 주어진 샘플에서 가장 잘 펩타이드를 제시할 가능성에 대응한다. 합성 펩타이드 검증은 오로지 $TSAs^{hi}$에 대해 수행되었다.

[0254]

[0255] MOI 식별의 강인성을 평가하기 위하여, 주어진 펩타이드의 관찰된 평균 체류 시간(RT)은 대량고속 MS로 식별된 MAP의 검증을 위한 2가지 동급 최상의 메트릭스와 상관시켰다: DeepLC 알고리즘에 의해 계산된 RT(Bouwmeester et al., 2020) 및 SSRcalc로 평가된 소수성 지수(Krokhin, 2006), 둘 다 펩타이드 서열에 기초하여 예측되었다. 이것은 비-정규 MOI의 RT 분포가 예측과 잘 상관되어 있고 정규 프로테옴-유래 펩타이드의 분포로부터 유의하게 상이하지 않았음(*F*-검정)을 나타내었으며, 이는 이들의 정확한 식별을 뒷받침한다(**도 3G**). 마지막으로, 모든 MS 데이터베이스 검색(처음에는 PEAKS 소프트웨어로 수행됨)은 Comet 알고리즘으로 반복되었다. 재식별의 백분율은 비-정규 MOI와 정규 펩타이드간에 유의한 차이를 나타내지 않았다(**도 3H**, 좌측 패널). MOI 중에서, 58개의 TSA^{hi} 중 52개(90%)가 재식별되었다(**도 3H**, 우측 패널). 두 개의 서로 다른 검색 엔진에 의해 식별된 MAP 간의 이러한 주요 중첩은 비-정규 MOI 식별의 강인성을 더욱 뒷받침한다.

[0256] **실시예 5: TSA^{hi}는 인트론의 번역으로부터 주로 유래되는 면역원성 MAP이다.**

[0257] 4가지 TSA-발견 접근법 결과의 조합은 총 47개의 HSA, 49개의 TAA, 36개의 TSA^{lo} 및 58개의 TSA^{hi}(중복 없이)를 산출하였다. 모든 MOI의 주된 특징은 **표 2**에 나열되어 있다. 정의에 따르면, TSA는 mTEC 및 정상 조혈 세포에서 뿐만 아니라 (GTEx로부터의) 모든 기관에서 역치 미만으로 발현되었다(**도 4A**). 중요하게는, 정상 조직에서의 TSA^{hi}-코딩 RNA의 발현은 표적을 벗어난 독성이 없는 임상 시험에서 앞서 사용된 TAA의 발현보다 체계적으로 열등하였다(Chapuis *et al.*, 2019; He *et al.*, 2020; Legat *et al.*, 2016; Qazilbash *et al.*, 2017). 이와 일관되게, 29개 비-악성 조직에서 식별된 인간 MAP를 함유하는 HLA 리간드 Atlas(https://www.biorxiv.org/content/10.1101/778944v1)에는 어떠한 TSA^{hi}도 존재하지 않는다. 이것은 표적화 TSA^{lo} 및 58개의 TSA^{hi}(중복 없이)의 안전을 뒷받침한다. TAA는 적어도 하나의 정상 조직에서 상승된 발현을 제시하는 한편 HSA 발현은 조혈 구획으로 제한되었다. AML 시편과 MPC 간의 FC의 비교는 TSA^{hi}가 TAA와 함께 최고 과발현(22배의 중앙값)을 제시하는 한편 HSA가 건강한 세포에서 최고 수준에서 발현된 것(0.6배의 중앙값)을 나타내었다(**도 4B**). 이들 결과는 모두 TSA^{hi}가 TSA의 특이성/안전성과 TAA의 과발현이라는 두 세계의 장점을 조

합하는 것을 나타낸다.

[0258] TSA는, 단지 13%가 정규 단백질 엑손으로부터 유래되었고 58%가 인트론으로부터 유래되었으므로, 대부분 게놈의 이른바 비-코딩 영역에서 유래된 것으로 식별되었다(**도 4C**). 단 하나도 AML 낮은 돌연변이 부담과 일치하는 돌연변이로부터 유래된 것은 단 하나도 없다(Lawrence *et al.*, 2013). TAA는 주로 단백질 코딩 엑손으로부터 유래된 반면, HSA 기원은 또한 조직-특이적 인트론 보유 및 ERE 발현 패턴을 보고한 이전 연구와 일치하여 비-코딩 영역에 의해 지배되었다(Middleton *et al.*, 2017; Larouche *et al.*, 2020). 8개의 TSA^{hi}가 정규 단백질-코딩 유전자로부터 유래되었지만, 이들은 안전한 TAA에 비해서 정상 조직에서 그들의 낮은 발현을 제공하는 것을 감안하면 안전한 표적으로서 간주될 수 있다(**도 4A**). 치료 표적으로서의 관련성을 뒷받침하기 위해, 이들 중 3개는 알려진 AML 바이오마커(*LTBP1*, *MYCN* 및 *PLPPR3*)로부터 유래되고, 나머지 5개는 미지의 기능을 갖거나 또는 증식, 분화 또는 약물 내성에 연루된다(**표 3**).

[0259] [표 3]

유도된 $TSAs^{hi}$가 식별된 정규 단백질-코딩 유전자의 특징

유전자 기호	전체 이름	앙상블 ID	암-관련 내용
LTBP1	잠재적인 전환 성장인자 베타 결합 단백질 1	ENSG00000049323	AML 에서 과발현되는, 잠재적인 TGF-β 의 분비 촉진(Wilson et al., 2006)
MYCN	MYCN 원종양유전자, bHLH 전사 인자	ENSG00000134323	AML 에서 과발현되는, 백혈병발생 기능(Liu et al., 2017; Wilson et al., 2006)
PLPPR3	인지질 포스파타제 관련 3	ENSG00000129951	AML 에서 과발현되는, ANPA32 백혈병 발생 유전자의 표적(Yang et al., 2018)
GCSAML	배중심 연관 신호전달 및 운동성 등	ENSG00000169224	증식과 연관하여, 이 유전자의 메틸화는 조혈성 악성 종양으로 변화된다(de Sá Machado Araújo et al., 2018)
DNAH10	다이네인 축세사 중쇄(Dynein axonemal heavy chain) 10	ENSG00000197653	증식 및 분화와 연관된, 원발성 섬모의 축세사 형성(Lagus et al., 2019)
ST8SIA6	ST8 알파-N-아세틸-뉴라미니드 알파-2,8-시알릴전달효소 6	ENSG00000148488	다제내성(Zhang et al., 2015)
LINC01835	C-형 렉틴 도메인 패밀리 4 멤버 O	ENSG00000267453	미지의 기능
MFSD2B	주요 촉진제 수퍼패밀리 도메인 함유 2B	ENSG00000205639	골수성 전구 세포에서의 추정적 후성 조절제(Johnson et al., 2015)

[0260]

[0261] TSA의 치료적 가치는 부분적으로는 환자가 공유하는 정도에 좌우된다. 1차 AML 간에 공유하는 TSA^{hi}를 평가하기 위하여, 437명의 환자에 대한 정제된 AML 모세포로부터의 RNA-seq 데이터를 포함하는 Leucegene 코호트(Lavallee et al., 2015; Macrae et al., 2013; Pabst et al., 2016)가 분석되었다. 대부분의 MAP는 상이한 HLA 알로타입에 의해 제시될 수 있으므로, 식별된 TSA^{hi}의 제시는 무차별 결합제를 고려해서 먼저 평가되었다. 유사한 에피토프를 제시하는 HLA 대립유전자와 함께 클러스터링하는 MHC클러스터 툴을 사용함으로써(Thomsen et al., 2013), 개별 TSA^{hi}를 제시 가능한 HLA 알로타입의 전체 세트가 외삽될 수 있었다(**표 4**). 이들 데이터에 기초하여, 세계 인구 중에서, 99.92%의 개인이 하나의 TSA^{hi}를 제시 가능한 1개 이상의 HLA-I 알로타입을 보유하는 것으로 나타났다. 다음에, 개별 TSA^{hi}는, TSA 코딩 전사체가 발현되고 환자가 이 TSA를 제시할 수 있었던 HLA 알로타입을 갖는 경우에만 주어진 AML 샘플에 존재한 것으로 간주되었다. 이러한 기준에 기초하여, Leucegene 코호트에서, 환자당 TSA^{hi}의 중앙값 수는 4이고, 93.6%의 환자가 적어도 1개의 TSA^{hi}를 제시한 것으로 예측될 수 있었다(**도 4F**).

[0262] [표 4]

MHC 클러스터에 의해 예측되는 바와 같은 유사한 펩타이드(난잡한 바인더)를 제시 가능한 HLA 대립유전자의 목록.

HLA 대립유전자	유사한 펩타이드(난잡한 바인더)를 제시 가능한 기타 대립유전자
HLA-A*01:01	
HLA-A*02:01	HLA-A*02:05 HLA-A*02:06 HLA-A*02:07
HLA-A*02:05	HLA-A*02:05 HLA-A*02:06 HLA-A*02:07
HLA-A*02:06	HLA-A*02:05 HLA-A*02:06 HLA-A*02:07
HLA-A*02:07	HLA-A*02:01 HLA-A*02:05 HLA-A*02:06
HLA-A*03:01	HLA-A*11:01

[0263]

HLA-A*11:01	HLA-A*03:01 HLA-A*31:01 HLA-A*68:01
HLA-A*23:01	HLA-A*24:02
HLA-A*24:02	HLA-A*23:01
HLA-A*25:01	HLA-A*26:01 HLA-A*66:01 HLA-B*15:02
HLA-A*26:01	HLA-A*25:01 HLA-A*66:01
HLA-A*29:02	HLA-A*30:02 HLA-B*15:02
HLA-A*30:01	
HLA-A*30:02	HLA-A*29:02
HLA-A*31:01	HLA-A*11:01 HLA-A*33:01 HLA-A*33:03 HLA-A*68:01
HLA-A*32:01	HLA-B*57:01 HLA-B*58:01
HLA-A*33:01	HLA-A*31:01 HLA-A*33:03 HLA-A*68:01
HLA-A*33:03	HLA-A*31:01 HLA-A*33:01 HLA-A*68:01
HLA-A*66:01	HLA-A*25:01 HLA-A*26:01
HLA-A*68:01	HLA-A*11:01 HLA-A*31:01 HLA-A*33:01 HLA-A*33:03
HLA-A*68:02	
HLA-B*07:02	HLA-B*35:02 HLA-B*35:03 HLA-B*55:01 HLA-B*56:01
HLA-B*08:01	
HLA-B*14:02	HLA-B*39:01
HLA-B*15:01	HLA-B*15:02 HLA-B*15:03 HLA-B*46:01
HLA-B*15:02	HLA-A*25:01 HLA-A*29:02 HLA-B*15:01 HLA-B*15:03 HLA-B*15:18 HLA-B*35:01 HLA-B*46:01
HLA-B*15:03	HLA-B*15:01 HLA-B*15:02 HLA-B*15:18
HLA-B*15:18	HLA-B*15:02 HLA-B*15:03
HLA-B*18:01	HLA-B*40:01 HLA-B*44:02 HLA-B*44:03 HLA-B*45:01
HLA-B*27:02	HLA-B*27:05 HLA-C*06:02 HLA-C*07:01
HLA-B*27:05	HLA-B*27:02
HLA-B*35:01	HLA-B*15:02 HLA-B*35:02 HLA-B*35:03 HLA-B*53:01
HLA-B*35:02	HLA-B*07:02 HLA-B*35:01 HLA-B*35:03 HLA-B*51:01 HLA-B*53:01 HLA-B*55:01 HLA-B*56:01
HLA-B*35:03	HLA-B*07:02 HLA-B*35:01 HLA-B*35:03 HLA-B*51:01 HLA-B*53:01 HLA-B*55:01 HLA-B*56:01
HLA-B*38:01	HLA-B*39:01
HLA-B*39:01	HLA-B*14:02 HLA-B*38:01
HLA-B*40:01	HLA-B*18:01 HLA-B*40:02 HLA-B*41:02 HLA-B*44:02 HLA-B*44:03 HLA-B*45:01
HLA-B*40:02	HLA-B*40:01 HLA-B*41:02 HLA-B*44:02 HLA-B*44:03 HLA-B*45:01
HLA-B*41:02	HLA-B*40:01 HLA-B*40:02 HLA-B*44:02 HLA-B*44:03 HLA-B*45:01
HLA-B*44:02	HLA-B*18:01 HLA-B*40:01 HLA-B*40:02 HLA-B*41:02 HLA-B*44:03 HLA-B*45:01
HLA-B*44:03	HLA-B*18:01 HLA-B*40:01 HLA-B*40:02 HLA-B*41:02 HLA-B*44:02 HLA-B*45:01
HLA-B*45:01	HLA-B*18:01 HLA-B*40:01 HLA-B*40:02 HLA-B*41:02 HLA-B*44:02 HLA-B*44:03
HLA-B*46:01	HLA-B*15:01 HLA-B*15:02 HLA-C*03:02 HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*12:02 HLA-C*12:03 HLA-C*16:01
HLA-B*51:01	HLA-B*35:02 HLA-B*35:03 HLA-B*52:01 HLA-B*53:01 HLA-B*55:01 HLA-B*56:01
HLA-B*52:01	HLA-B*51:01
HLA-B*53:01	HLA-B*35:01 HLA-B*35:02 HLA-B*35:03 HLA-B*51:01
HLA-B*55:01	HLA-B*07:02 HLA-B*35:02 HLA-B*35:03 HLA-B*51:01 HLA-B*56:01
HLA-B*56:01	HLA-B*07:02 HLA-B*35:02 HLA-B*35:03 HLA-B*51:01 HLA-B*55:01
HLA-B*57:01	HLA-A*32:01 HLA-B*58:01
HLA-B*58:01	HLA-A*32:01 HLA-B*57:01
HLA-C*03:02	HLA-B*46:01 HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*12:02 HLA-C*12:03 HLA-C*16:01
HLA-C*03:03	HLA-B*46:01 HLA-C*03:02 HLA-C*03:04 HLA-C*08:01 HLA-C*08:02 HLA-C*12:02 HLA-C*12:03 HLA-C*15:02 HLA-C*16:01

[0264]

[0265]

HLA-C*03:04	HLA-B*46:01 HLA-C*03:02 HLA-C*03:04 HLA-C*08:01 HLA-C*08:02 HLA-C*12:02 HLA-C*12:03 HLA-C*15:02 HLA-C*16:01
HLA-C*04:01	HLA-C*07:02 HLA-C*14:02
HLA-C*05:01	HLA-C*08:01 HLA-C*08:02
HLA-C*06:02	HLA-B*27:02 HLA-C*07:01 HLA-C*07:02
HLA-C*07:01	HLA-B*27:02 HLA-C*06:02 HLA-C*07:02 HLA-C*14:02
HLA-C*07:02	HLA-C*04:01 HLA-C*06:02 HLA-C*07:01 HLA-C*14:02
HLA-C*08:01	HLA-B*46:01 HLA-C*03:02 HLA-C*03:03 HLA-C*03:04 HLA-C*05:01 HLA-C*08:02 HLA-C*12:02 HLA-C*12:03 HLA-C*15:02 HLA-C*16:01
HLA-C*08:02	HLA-C*03:03 HLA-C*03:04 HLA-C*05:01 HLA-C*08:01 HLA-C*15:02
HLA-C*12:02	HLA-B*46:01 HLA-C*03:02 HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*12:03 HLA-C*15:02 HLA-C*16:01
HLA-C*12:03	HLA-B*46:01 HLA-C*03:02 HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*12:03 HLA-C*15:02 HLA-C*16:01
HLA-C*14:02	HLA-C*04:01 HLA-C*07:01 HLA-C*07:02 HLA-C*16:01
HLA-C*15:02	HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*08:02 HLA-C*12:02 HLA-C*12:03
HLA-C*16:01	HLA-B*46:01 HLA-C*03:02 HLA-C*03:03 HLA-C*03:04 HLA-C*08:01 HLA-C*12:02 HLA-C*12:03 HLA-C*14:02

[0266] 초기 진단 시 대 재발 시 분석된 AML 샘플(비정합 샘플)에서 TSA^{hi}의 수를 비교할 때, 두 군 사이에 차이가 발견되지 않았다(**도 4G**). 또 다른 연구에서 진단 시와 동종 조혈 세포 이식 후 재발 시 얻어진 AML 모세포의 정합된 샘플에서 환자의 HLA 대립유전자가 제시할 수 있는 TSA^{hi}의 RNA 발현을 비교할 때에도 차이가 발견되지 않았다(Toffalori et al., 2019)(**도 4H**). As 백혈병 줄기세포(LSC)가 재발의 주요 매개체이기 때문에(Shlush et al., 2017), TSA^{hi} 및 HLA RNA 발현이 또한 또 다른 연구에서 얻어진 분류된 LSC 대 모세포 RNA-seq 데이터에서 평가되었고(Corces et al., 2016), 두 세포 집단 간의 차이는 발견되지 않았다(**도 4I 내지 J**). 그럼에도 불구하고, 유전자 세트 농축 분석(gene set enrichment analysis: GSEA)을 사용함으로써, 높은 수의 TSA^{hi}를 발현하는 환자가 또한 잘 확립된 LSC 유전자 시그니처의 높은 수준을 발현한 것으로 밝혀졌다(Eppert *et al.*, 2011)(**도 4K**). 종합하면, 이들 결과는 TSA^{hi}의 높은 면역원성을 더욱 뒷받침하고, 이들은 진단 시 또는 재발 시 거의 모든 AML 환자에서 표적화될 수 있었음을 입증하고 있다. 따라서, TSA^{hi}의 면역 표적화가 질환의 모든 병기에서 상정될 수 있었고 LSC를 제거할 가능성을 가질 것이라고 결론 내릴 수 있다.

[0267] **실시예 6: 많은 TSA^{hi}의 제시는 더 양호한 생존율과 상관관계가 있다.**

[0268] 다음에, 환자 생존에 대한 진단 시 TSA^{hi} 제시의 영향이 조사되었다. 놀랍게도, 가장 높은 수(상위 사분위수)의 TSA^{hi}를 발현하는 환자는 나머지 코호트보다 유의하게 더 양호한 생존율을 제시하였다(**도 5A**). 다수 TSA^{hi}의 제시와 관련된 생존 이점은, 연령, 세포유전학 위험, *NPM1* 및 *FLT3-ITD* 돌연변이와 같은 다른 공지된 예후 인자와 함께, 다변량 분석에서 유의하게 유지되었다(**도 5B**). 중요하게는, HLA $_{pred}$제시와는 독립적으로 수행된 동일한 비교는 높은 발현자와 낮은 발현자 간에 차이가 없는 것을 나타냈는데(**도 11A, B**), 이는 TSA^{hi}의 보호 효과가 HLA-제한적이었음을 의미한다. TAA, HSA 또는 TSA^{lo}에 대해 수행된 동일한 분석은 생존에 대한 유의한 영향을 나타내지 않았다(**도 11C 내지 H**). 이들 데이터는 TSA^{hi}가 자발적인 항-AML 면역 반응을 유도하는데 충분하게 면역원성임을 시사한다.

[0269] TSA^{hi}에 의해 제공되는 생존 이점이 누적 HLA $_{pred}$제시로부터 초래된 것을 입증하기 위하여, 높은 대 낮은 TSA^{hi} 환자의 로그-순위 p-값은 분석(1000개의 랜덤 순열/수)으로부터 증가하는 수의 TSA^{hi}(58개 중 1 내지 29개)의 랜덤 제거 후 계산되었다. 증가하는 수의 TSA^{hi}의 확률적 제거는 낮은 발현자(다른 모든 환자)에 비해서 높은 발현자(HLA-TSA^{hi} 계수치의 상위 사분위수)에 대한 상당한 생존 이점의 손실을 빠르게 초래하였다(**도 5C 내지 D**). 개별 TSA^{hi}의 차감에 따른 생존 이점의 점진적 감소는 대다수의 TSA^{hi}가 이 생존 이점에 기여함을 시사한다. 더 많은 비율의 환자에 의해 제시된 TSA^{hi}는 p-값에 가장 큰 영향을 미쳤다(**도 5E**). 마찬가지로, 로그-순위 분석으로부터 (5% 초과의 환자에 의해 공유된) 공통 HLA 대립유전자의 제거는 낮은 빈도 대립유전자의 제거보다 p-값에 대한 더 많은 영향을 가졌다(**도 5F**). 종합하면, 이들 데이터는 환자 생존율에 대한 TSA^{hi} 제시의 HLA-제한된 이

점을 입증한다. 다음 실험에서, TSA^{hi} 제시와 연계된 생존 이점에 대한 가장 간결한 설명이 조사된다: TSA^{hi}는 자발적 보호 항-AML 면역 반응을 유도한다.

[0270] **실시예 7: TSA^{hi} 제시는 세포독성 T 세포 반응을 촉발시킨다**

[0271] TSA^{hi} 및 기타 MOI의 면역원성(즉, 면역 반응을 유도하는 능력)의 평가의 서곡으로서, T-세포 반응의 가능성을 예측하기 위해 공개 TCR 데이터베이스에 의존하는 기계 학습 알고리즘인 Repitope(Ogishi and Yotsuyanagi, 2019)가 사용되었다. 흉선 상피 세포에 의해 제시되거나(Adamopoulou et al., 2013) 또는 각각 음성 및 양성 대조군으로서 HIV로부터 유래된 MAP를 사용한 경우, Repitope 예측은 TAA가 대부분 비-면역원성인 한편 MOI의 3개의 다른 군이 HIV 펩타이드만큼 면역원성인 것을 나타내었다(**도 6A**). 따라서, TAA는, 3개의 다른 군에 대해서 그리고 IEDB에서 면역원성으로서 보고된 1411개의 MAP의 세트에 대해서 mTEC(대략 12.1 rphm)의 높은 발현을 제시하였다(**도 6B**). 비-TAA MOI는 모두 심지어 다른 면역원성 펩타이드에 비해서 mTEC의 매우 낮은 RNA 발현을 제시하였는데, 이는 이들의 면역원성을 뒷받침한다. Repitope 예측을 검증하기 위하여, 시험관내 T-세포 검정은 가장 높은 면역원성인 것으로 예측된 HLA-A*02:01-제시된 TSA^{hi}, 즉, ALPVALPSL에서 시작하여 수행되었다. IFN-γ ELISpot에서의 양성 대조군으로서, E*L*AGIGILTV 에피토프는, 가장 면역원성인 인간 MAP 중 하나이기 때문에 사용되었다(Dutoit et al., 2002; Hesnard et al., 2016). ALPVALPSL의 면역원성은 E*L*AGIGILTV의 것과 유사하였다(**도 6C**). 2개의 다른 유망한 TSA^{hi}의 ELISpot은 또한 그들의 면역원성을 뒷받침하였다(**도 6D**). 이들 TSA^{hi}의 경우, 사이토카인 분비 검정 및 덱스트라머 염색이 또한 수행되는데, 이는 ELISpot 결과를 확인하고 면역 반응의 특이성을 뒷받침하였다(**도 6E 내지 F**). TSA^{hi}가 자발적이고도 특이적인 T-세포 클론형 확장을 유도할 수 있는 것을 더욱 입증하기 위하여, 상이한 풀의 TSA^{hi}로 자극되는 말초 혈액 T 세포의 단기 배양액이 TCR 시퀀싱을 통해서 분석되는(Danilova et al., 2018) 특이적 T 세포의 기능적 확장(FEST) 검정이 수행되었다. 5개의 시험된 TSA^{hi}의 각각의 풀은 9 내지 10개의 상이한 클론형의 특이적 확장을 유도하였는데, 이는 이들의 자발적 면역원성을 뒷받침하였다(**도 6G** 및 **표 5**).

[0272] [표 5]

TSAs^{hi}의 풀에 반응하는 특정 T 세포의 기능 확장(FEST) 검정.

풀 #	풀에서의 TSAs^{hi}	유의하게 확장된 클론형(서열번호)	FDR	오즈비
1	SLLSGLLRA ALPVALPSL ALDPLLLRI IASPIALL SLDLLPLSI	CSARGDREYEQYF (191)	4.84E-52	Inf
1		CASTVVAGNSSPLHF (192)	1.2E-07	18.829
1		CASSQDGIWGAYEQYF (193)	1.95E-07	Inf
1		CASSVDAGGNYEQYF (194)	0.000665	12.293
1		CASSLGGQGLSYGYTF (195)	0.011005	Inf
1		CASSYRPNEQYF (196)	0.011005	Inf
1		CASSGTDLNQPQHF (197)	0.032852	Inf
1		CASSSNFEPLHF (198)	0.032852	Inf
1		CASSWGGSNTGELFF (199)	0.032852	Inf
2	LTDRIYLTL VLFGGKVSGA LGISLTLKY FNVALNARY TLNQGINVYI	CASSEYRALNTEAFF (200)	4.67E-30	59.789
2		CASSESGTGGQPQHF (201)	1.14E-08	3.455
2		CASSRTGENTEAFF (202)	0.003813	3.352
2		CASSSTDRQHYGYTF (203)	0.007052	3.764
2		CASSDRTGGSSNEKLFF (204)	0.008031	8.213
2		CATSRSGDSNQPQHF (205)	0.008031	8.213
2		CASSYVLNTEAFF (206)	0.010284	10.837
2		CASRESGQMNEKLFF (207)	0.015395	Inf
2		CASSDGGEGTYGYTF (208)	0.028405	Inf
2		CASTSWTGFGPNYGYTF (209)	0.028405	Inf
3	LRSQILSY KILDVNLRI HSLISIVYL KLQDKEIGL AQDIILQAV	CASSADSSLGGYTF (210)	1.13E-10	45.434
3		CSARDLAGGTYEQYF (211)	8.08E-10	Inf
3		CASSPPTGEYEKLFF (212)	9.54E-09	Inf
3		CARSFGGFF (213)	3.78E-06	Inf
3		CASSLPSGILYEQYF (214)	3.78E-06	Inf
3		CASAPGGMPYGYTF (215)	1.16E-05	Inf
3		CASSLQDTDYNEQFF (216)	0.005265	Inf
3		CASSDPGTSGVFTGELFF (217)	0.009073	Inf
3		CSARLGTGELFF (218)	0.016806	Inf
3		CASSLGRGYETQYF (219)	0.019961	3.55

HLA-A*02:01, -A*29:02, -B*15:01, -B27:05, -C*01:02, -C*03:04 건강한 공여체의 HLA 대립유전자에 의해 제시될 수 있었던 모든 TSAs^{hi}. TCR-seq 는 Adaptive Biotechnologies 사에 의해 제작되었고, 원시 데이터는 FEST 분석 툴 http://www.stat-apps.onc.jhmi.edu/FEST 로 처리되었다. 여기에서는 풀의 번호, 각 풀에 존재하는 TSAs^{hi}, 각 풀에서 유의하게 확장된 각 클론형의 서열, 및 FEST 분석 툴에 의해 제공된 FDR 및 오즈비(Odds ratio)가 보고되어 있다.

[0273]

[0274] 다음에, "생체내 진리(*in vivo veritas*)" 때문에, 437명의 Leucegene 환자로부터의 전사체 데이터에 대한 심층 분석은 T 세포에 의한 TSA^{hi}의 잠재적인 생체내 인식을 평가하기 위하여 수행되었다. 우선, T 세포의 TCR 레퍼토리 다양성은 TRUST4 알고리즘으로 평가하였다(Zhang et al., 2019). (본 명세서에서 비-면역원성 대조군으로서 사용된) TAA와는 대조적으로, 상승된 TSA^{hi} 수의 $_{pred}$제시는 저감된 TCR 레퍼토리 다양성과 연관되었는데, 이는 항-TSA^{hi} 클론형의 확장을 시사한다(**도 6H**). 이 확장의 특이성을 입증하기 위하여, ERGO 알고리즘이 MOI-TCR 상호작용을 예측하는 데 사용되었다(Springer et al., 2020). 항-MOI 클론형을 식별하기 위한 ERGO 가능성>80%에서, 높은 수의 TSA^{hi}를 갖는 환자는 또한 모든 검출된 CDR3 중에서 더 큰 빈도의 항-TSA^{hi} 클론형을 가졌다(**도 6I**). TAA에 대해서는 유사한 상관관계가 나타나지 않았다(**도 6J**). 다음에, 각각의 AML 샘플에 의해 $_{pred}$제시된 MOI를 인식 가능한 항-MOI 클론형의 비율(즉, 동족 TCR-MOI 상호작용의 빈도)이 계산되었다. 이 비율은 $_{pred}$제시된 MOI에 따라서 정규화되었는데, 그렇지 않았다면 더 많은 MOI의 제시는 더 높은 비율의 항-$_{pred}$제시된 MOI 클론형의 검출을 초래할 것이기 때문이다. 이것은 TSA^{hi} $_{pred}$제시가 TAA보다 극적으로 더 높은 빈도의 특정 T-세포 인식과 연관되었음을 나타내었다(**도 6K**).

[0275] 항-TSA^{hi} T-세포 인식에 비추어, TSA^{hi} $_{pred}$제시가 활성화된 CD8 T 세포의 침윤과 연관되어야 한다고 추론되었다. 흥미롭게도, TSA^{hi} 전사체의 다양성은 AML 샘플에서 *CD8A*+*CD8B* 발현과 역상관관계에 있는 한편, 이들의 $_{pred}$제시의 다양성은 그렇지 않았다(**도 6L 내지 M**). 이것은 TSA^{hi} 전사체의 높은 다양성이 (TSA에 대해서 예상될 수 있었

던 것처럼) AML 샘플에서 약간 더 높은 모세포 순도를 반영했음을 시사하였다. 이 가능한 편향을 피하기 위해 그리고 HLA-TSA^{hi}의 수가 발현된 TSA^{hi}의 수와 수학적으로 연계되기 때문에, $_{pred}$제시된 TSA^{hi}의 수는 발현된 TSA^{hi} 전사체의 수로 정규화되었고, 유전자 발현량 차이는 정규화된 $_{pred}$제시가 중앙값보다 높았던 환자 대 중앙값보다 낮았던 환자에서 분석되었다. 놀랍게도, TSA^{hi} $_{pred}$제시와 양성적으로 연관된 123개의 유전자 중에서, *CD8A*, *CD8B*, *GZMA*, *GZMB*, *IL2RB*, *PRF1* 및 *ZAP70*을 비롯하여 몇 가지는 T-세포 활성화 및 세포용해와 연관되었다(**도 6N**). 특히, 이들 123개 유전자와 연관된 GO 용어는 T-세포 활성화 및 분화와 배타적으로 관련되었다(**도 6O**). *CD4* 유전자는 차등적으로 발현되지 않았고, GO 용어는 하향조절된 유전자와 유의하게 연관될 수 없었다. 그러므로, TSA^{hi} $_{pred}$제시는 활성화된 CD8 T 세포의 더 높은 풍부도와 연관된다고 결론을 내렸다.

[0276] **실시예 8: TSA^{hi} RNA 발현은 면역편집, AML 드라이버 돌연변이 및 후성적 비정상 징후와 연관된다**

[0277] TSA^{hi}의 잠재적인 치료 가치를 감안할 때, 이들의 생물발생에 대한 통찰력을 얻는 것이 바람직하다. 이 분석을 위하여, 고도로 발현된 TSA^{hi}의 계수치(HE-TSA^{hi}), 즉, 주어진 TSA^{hi}의 비-null 발현을 갖는 모든 환자에 걸쳐서 이들의 중앙 발현보다 더 높은 수준에서 발현된 TSA^{hi}의 계수치가 각 Leucegene 환자에 기인되었다. 이어서, 특정 유전자의 발현이 각 단백질 코딩 유전자의 발현과 HE-TSA^{hi} 계수치 간의 쌍별 피어슨 상관을 수행함으로써 TSA^{hi} 발현에 연결될 수 있었는지를 평가하였다. 이것은 MAP 제시에 연루된 유전자(*HLA-A*, *HLA-B*, *HLA-C*, *B2M* 및 *NLRC5*)의 발현과 HE-TSA^{hi}의 수 간의 일관된 역 상관관계를 나타내었는데, 이는 상승된 TSA^{hi} 발현에 반응한 면역편집의 발현을 시사한다(**도 7A** 및 **도 12A-B**). 면역편집은 또한 수지상 세포 식세포작용의 저해와 연루된 면역관문 분자인 *CD47*(Majeti *et al.*, 2009) 및 백혈병 세포에 의한 PD-L1 발현을 촉진시키는 *CD84*(Lewinsky *et al.*, 2018)와의 양성 상관관계에 의해 뒷받침되었다. *NPM1* 돌연변이가 PD-L1(CD274) 발현을 조절할 수 있기 때문에(Greiner *et al.*, 2017), *NPM1*^{mut} 및 *NPM1*^{wt} AML 환자는 별도로 분석되었다. 이 분석은 중앙값보다 높은 HE-TSA^{hi} 계수치를 갖는 *NPM1*^{wt} 환자가 열등한 HE-TSA^{hi} 계수치를 갖는 환자보다 유의하게 더 높은 수준의 PD-L1을 발현한 것을 나타내었다(**도 7B**).

[0278] 다음에, HE-TSA 계수치와 상관관계가 있는 유전자 경로가 분석되었다(**도 7C** 및 **표 6**). 음성적으로 상관된 경로는 세포 증식(수송 및 세포 조직화도 포함), 미토콘드리아 OXPHOS 및 프로테아좀-매개 단백질 이화작용과 연루된 생물학적 과정을 포함하였다. 흥미롭게도, 미토콘드리아 활성의 저해는 MHC-I 발현을 저감시키는 것으로 나타났고, 암세포에 의한 면역회피기전으로서 사용될 수 있었다(Charni *et al.*, 2010). 마찬가지로, 단백질 분해의 저해는 MHC-I 분자에 의한 제시를 위한 더 적은 양의 펩타이드로 이어질 수 있고, 따라서 TSA^{hi}를 제시할 가능성을 저감시킬 것이다(Tripathi *et al.*, 2016). 마지막으로, 유사분열-관련 과정의 저감은 MHC-I 하향조절의 부작용일 수 있는데, 그 이유는 두 과정이 모두 *NLRC5*에 의해 조절되기 때문이다(Wang *et al.*, 2019). 종합하면, 이들 데이터는 TSA^{hi} 발현이 AML 세포에 대한 면역편집 기전으로서 작용할 수 있는 다양한 반응과 연계되는 것을 나타낸다.

[0279] 음성적으로 상관된 경로와 대조적으로, 양성적으로 상관된 경로는 조절 과정으로 제한되었다(**도 7D**). 따라서, 양성적 상관 유전자의 16.1%(대 2.5%의 음성적 상관)는 전사 인자였는데, 이는 TSA^{hi}의 전사를 직접 담당할 수 있었다. 이들 중, 최상의 상관된 유전자는, 게놈 임프린팅, 즉, DNA 메틸화에 연계된 후성적 과정(Takahashi *et al.*, 2019)의 조정자인 *ZNF445*였다(**도 7A**). *ZNF445* 기능은 DNA 메틸화에 의존하기 때문에, TSA^{hi} 발현과 전형적으로 DNA 메틸화 비정상에 연계된 AML 돌연변이 간의 가능한 연관이 조사되었다. 3개의 가장 빈번한 AML 드라이버 돌연변이(*NPM1*^{mut}, *FLT3-ITD* 및 *DNMT3A*^{mut})가 먼저 시험되었고, 3개 모두 높은 HE-TSA^{hi} 계수치를 발현하는(중앙값보다 높은) 환자에서 유의하게 농축된 것으로 발견되었다(**도 7E**). 또한, 2개 또는 3개의 동반 돌연변이를 제시하는 환자는 하나 또는 전혀 제시하지 않는 환자보다 높은 수의 HE-TSA^{hi}를 지녔다(**도 7F**). MS 분석에서 사용된 19개의 AML 시편 중 12개가 *FLT3*-ITD 또는 *NPM1* 돌연변이를 나타냈다. 다른 빈번한 AML 돌연변이와 관련하여, *IDH2* 및 이중대립유전자 *CEBPA* 돌연변이도 상승된 HE-TSA^{hi} 계수치와 양성적으로 연관된 한편 *ASXL1*, *SRSF2* 및 *U2AF1* 돌연변이는 음성적으로 연관되었고, *FLT3-TKD*, *IDH1*, *RUNX1*, *TET2*, *TP53* 및 *WT1*은 연관되지 않은 것

으로 밝혀졌다(**도 12D**). *NPM1*, *DNMT3A*, *IDH2* 및 *CEBPA*^{bi} 돌연변이가 비정상적인 메틸화 프로파일과 연관되므로(Figueroa *et al*., 2010a; Figueroa *et al*., 2010b; Ley *et al*., 2013), 상승된 HE-TSA^{hi} 계수치를 갖는 이들의 상관관계는 TSA^{hi} 발현에서 후성적 조절이상의 암시를 뒷받침한다.

[0280] [표 6]

Leucegene 코호트에서 HE-TSAshi 계수치와 양성적으로 또는 음성적으로 상관이 있는 GO 용어의 목록.

양성적/음성적 상관?	GO 용어 ID	p-값	질의 목록에서의 적중 계수치	GO 용어에서의 총 유전자	명칭
음성적	278	5,63E-19	74	381	유사분열 세포주기
음성적	9987	2,48E-13	599	9339	세포 과정
음성적	51443	7,60E-13	27	71	유비퀴틴-단백질 리가제 활성의 양성 조절
음성적	280	2,14E-12	48	232	핵 분열
음성적	7067	2,14E-12	48	232	유사분열
음성적	7049	2,48E-12	100	795	세포주기
음성적	51351	2,61E-12	27	74	리가제 활성의 양성 조절
음성적	22402	2,72E-12	82	583	세포주기 과정
음성적	6091	5,20E-12	56	312	전구체 대사산물 및 에너지의 생성
음성적	87	7,37E-12	48	239	유사분열 세포주기의 M 상
음성적	51437	7,83E-12	25	65	유사분열 세포주기에 연루된 유비퀴틴-단백질 리가제 활성의 양성 조절
음성적	48285	1,04E-11	48	241	세포기관 분열
음성적	6996	2,30E-11	142	1377	세포기관 조직
음성적	51301	2,86E-11	55	314	세포 분열
음성적	51439	4,07E-11	25	69	유사분열 세포주기에 연루된 유비퀴틴-단백질 리가제 활성의 조절
음성적	51438	5,09E-11	27	82	유비퀴틴-단백질 리가제 활성의 조절
음성적	51340	1,40E-10	27	85	리가제 활성의 조절
음성적	51436	4,26E-10	23	63	유사분열 세포주기에 연루된 유비퀴틴-단백질 리가제 활성의 음성 조절
음성적	31398	5,15E-10	28	96	단백질 유비퀴틴화의 양성 조절
음성적	31145	6,32E-10	23	64	후기-촉진 복합체-의존적 프로테아좀 유비퀴틴-의존적 단백질 이화과정
음성적	51352	1,97E-09	23	67	리가제 활성의 음성 조절

[0281]

[0282]

음성적	51444	1,97E-09	23	67	유비퀴틴-단백질 리가제 활성의 음성 조절
음성적	43161	2,40E-09	33	140	프로테아좀 유비퀴틴-의존적 단백질 이화과정
음성적	10498	2,40E-09	33	140	프로테아좀 단백질 이화과정
음성적	31396	9,12E-09	30	122	단백질 유비퀴틴화의 조절
음성적	31397	1,24E-08	24	79	단백질 유비퀴틴화의 음성 조절
음성적	16043	1,47E-08	212	2557	세포 성분 조직화
음성적	279	4,28E-08	53	351	M 상
음성적	31400	4,99E-08	31	138	단백질 변형 과정의 음성 조절
음성적	44237	5,13E-08	353	4974	세포 대사과정
음성적	8152	1,63E-07	404	5938	대사과정
음성적	22403	2,44E-07	59	435	세포주기상
음성적	8104	3,87E-07	97	921	단백질 국소화
음성적	45184	5,51E-07	85	767	단백질 국소화의 확립
음성적	15031	6,33E-07	84	756	단백질 수송
음성적	32269	8,96E-07	36	200	세포 단백질 대사과정의 음성 조절
음성적	51248	1,42E-06	37	213	단백질 대사과정의 음성 조절
음성적	33036	3,32E-06	108	1111	거대분자 국소화
음성적	6119	4,67E-06	24	102	산화적 인산화
음성적	31401	4,68E-06	38	232	단백질 변형 과정의 양성 조절
음성적	51179	8,10E-06	227	2978	국소화
음성적	16192	9,74E-06	68	589	소포-매개 수송
음성적	9056	9,96E-06	99	1005	이화과정
음성적	22900	1,21E-05	25	115	전자 수송 사슬
음성적	32270	3,08E-05	42	291	세포 단백질 대사과정의 양성 조절
음성적	51234	4,69E-05	201	2606	국소화의 확립
음성적	31399	5,75E-05	48	366	단백질 변형 과정의 조절
음성적	6511	7,23E-05	40	277	유비퀴틴-의존적 단백질 이화과정
음성적	7059	8,48E-05	20	83	염색체 분리
음성적	43632	1,21E-04	40	282	변형-의존적 거대분자 이화과정
음성적	19941	1,21E-04	40	282	변형-의존적 단백질 이화과정
음성적	44092	1,50E-04	48	377	분자 기능의 음성 조절
음성적	51247	1,69E-04	42	308	단백질 대사과정의 양성 조절
음성적	9057	1,73E-04	58	502	거대분자 이화과정
음성적	16044	2,11E-04	48	381	세포막 조직화
음성적	51246	2,15E-04	68	635	단백질 대사과정의 조절
음성적	51603	2,24E-04	42	311	세포 단백질 이화과정에 연루된 단백질분해
음성적	61024	2,29E-04	48	382	막 조직화
음성적	32268	2,56E-04	62	559	세포 단백질 대사과정의 조절
음성적	44257	2,97E-04	42	314	세포 단백질 이화과정
음성적	30163	3,43E-04	44	339	단백질 이화과정
음성적	51641	3,79E-04	92	977	세포 국소화

음성적	44265	4,64E-04	52	440	세포거대분자 이화과정
음성적	46907	5,91E-04	69	665	세포내 수송
음성적	34613	7,00E-04	51	433	세포 단백질 국소화
음성적	51649	8,38E-04	83	865	세포에서의 국소화의 확립
음성적	70727	8,74E-04	51	436	세포거대분자 국소화
음성적	6810	9,25E-04	193	2572	수송
음성적	45333	1,17E-03	20	96	세포 호흡
음성적	44248	1,28E-03	76	775	세포 이화과정
음성적	42773	1,78E-03	15	57	ATP 합성 결합된 전자 수송
음성적	42775	1,78E-03	15	57	미토콘드리아 ATP 합성 결합된 전자 수송
음성적	22904	1,91E-03	16	65	호흡성 전자 수송 사슬
음성적	55114	2,07E-03	66	646	산화 환원
음성적	43086	6,03E-03	38	301	촉매 활성의 음성 조절
음성적	15980	6,23E-03	24	145	유기 화합물의 산화에 의한 에너지 유도
음성적	7010	1,18E-02	49	448	세포골격 조직화
음성적	6886	1,75E-02	42	364	세포내 단백질 수송
음성적	226	2,27E-02	23	145	미세소관 세포골격 조직화
음성적	7017	2,67E-02	32	247	미세소관-기반 과정
음성적	44093	2,95E-02	62	640	분자 기능의 양성 조절
음성적	16052	3,12E-02	19	107	탄수화물 이화과정
음성적	33043	4,83E-02	31	242	세포기관 조직화의 조절
음성적	48193	5,93E-02	21	132	골지 소포 수송
음성적	6508	7,37E-02	67	730	단백질분해
음성적	65009	9,69E-02	89	1063	분자 기능의 조절
음성적	43085	1,04E-01	54	553	촉매 활성의 양성 조절
음성적	7033	1,34E-01	11	44	액포 조직화
음성적	7264	1,39E-01	34	292	작은 GTP 가수분해효소 매개 신호 전달
음성적	6007	1,64E-01	12	53	글루코스 이화과정
음성적	44238	1,93E-01	334	5270	1 차 대사과정
음성적	16050	2,44E-01	12	55	소포 조직화
음성적	7040	2,93E-01	8	25	리소좀 조직화
음성적	70585	2,97E-01	9	32	미토콘드리아에서의 단백질 국소화
음성적	6626	2,97E-01	9	32	미토콘드리아에 대한 단백질 표적화
음성적	910	3,26E-01	11	48	세포질 분열
음성적	46365	3,60E-01	13	66	단당류 이화과정
음성적	6122	3,78E-01	4	5	미토콘드리아 전자 수송, 유비퀴놀에서 사이토크롬 c 로
음성적	51128	4,05E-01	51	538	세포 성분 조직화의 조절
음성적	44267	5,10E-01	153	2146	세포 단백질 대사과정
음성적	6120	6,50E-01	10	43	미토콘드리아 전자 수송, NADH 에서 유비퀴논으로
음성적	8064	7,23E-01	12	61	액틴 중합 또는 해중합의 조절
음성적	50790	7,39E-01	75	906	촉매 활성의 조절
음성적	51656	8,00E-01	13	71	세포기관 국소화의 확립
음성적	70	8,23E-01	9	36	유사분열 자매 염색분체 분리

[0283]

음성적	44275	8,29E-01	14	81	세포 탄수화물 이화과정
음성적	19320	1,01E+00	12	63	핵소스 이화과정
음성적	30832	1,01E+00	12	63	액틴 필라멘트 길이의 조절
음성적	819	1,04E+00	9	37	자매 염색분체 분리
음성적	19538	1,13E+00	178	2604	단백질 대사과정
음성적	10324	1,18E+00	26	221	막 함입
음성적	6897	1,18E+00	26	221	엔도사이토시스
음성적	32956	1,20E+00	15	94	액틴 세포골격 조직화의 조절
음성적	43254	1,20E+00	15	94	단백질 복합체 조립의 조절
음성적	30833	1,23E+00	11	55	액틴 필라멘트 중합의 조절
음성적	10564	1,25E+00	19	138	세포주기 과정의 조절
음성적	7346	1,36E+00	22	174	유사분열 세포주기의 조절
음성적	60316	1,49E+00	3	3	리아노딘-민감성 칼슘-방출 통로 활성의 양성 조절
음성적	51640	1,53E+00	15	96	세포기관 국소화
음성적	32271	1,60E+00	12	66	단백질 중합의 조절
음성적	44283	1,92E+00	42	443	소분자 생합성 과정
음성적	32970	1,94E+00	15	98	액틴 필라멘트-기반 과정의 조절
양성적	45449	1,70E-14	184	2621	전사의 조절
양성적	19219	6,15E-14	201	3013	핵염기, 뉴클레오사이드, 뉴클레오타이드 및 핵산 대사과정의 조절
양성적	51171	7,28E-14	202	3039	질소 화합물 대사과정의 조절
양성적	10468	6,82E-13	194	2926	유전자 발현의 조절
양성적	10556	1,17E-12	191	2876	거대분자 생합성 과정의 조절
양성적	31326	4,39E-12	196	3020	세포 생합성 과정의 조절
양성적	9889	1,03E-11	196	3044	생합성 과정의 조절
양성적	80090	4,31E-11	217	3553	1 차 대사과정의 조절
양성적	31323	1,80E-10	223	3735	세포 대사과정의 조절
양성적	60255	1,97E-10	207	3375	거대분자 대사과정의 조절
양성적	19222	6,73E-10	229	3917	대사과정의 조절
양성적	51252	1,46E-08	130	1855	RNA 대사과정의 조절
양성적	6355	2,28E-08	127	1806	DNA-의존적인 전사의 조절
양성적	50909	5,95E-04	12	44	맛의 감각 지각

[0284]

[0285] 마지막으로, TSA^{hi} 발현이 AML 환자, 예컨대, 프랑스-미국-영국(FAB) 유형(**도 12E 내지 H**)에서 이의 존재를 예측할 수 있게 하는 다른 임상적 특징부와 연결될 수 있었는지를 조사하였다. 놀랍게도, 높은 계수치의 HE-TSA^{hi}를 발현하는 환자는 각각 M1 및 M5 AML에서 과다 표시 및 과소 표시되었다. 따라서, 성숙하지 않은 AML 및 정상 핵형을 갖는 환자는 최고 수준의 TSA^{hi}를 제시하였다. 따라서, 이것은 TSA를 발견하는 데 사용된 샘플에서 FAB M1 AML의 과다 표시에 기인되는 것으로 가정되었고(19개 중 9개), TSA^{hi}의 대부분이 인트론 영역(37/58, 인트론에 위치된 ERE-유래 TSA^{hi}를 포함)에 위치되므로, 이것은 FAB 유형들 간의 상이한 패턴의 인트론 보유의 존재에 의해 설명될 수 있었다. 따라서, AML에 특이적으로 보유된 인트론 상에서 수행된 무감독 공통 클러스터링은 이들의 FAB 유형에 따른 환자의 명백한 클러스터링을 나타내었다(**도 7G**). 종합하면, 이들 데이터는 TSA^{hi} 발현이 AML 아형에 특이적인 인트론 보유 패턴과 연결되는 것을 나타낸다.

[0286] 본 발명이 이의 구체적인 실시형태에 의해서 위에서 설명되었지만, 첨부된 청구범위에 정의된 바와 같은 본 발명의 정신 및 특성을 벗어나는 일 없이 변형될 수 있다. 청구범위에서, 단어 "포함하는"(comprising)은 "를 포함하지만, 이로 제한되지 않는"이라는 어구에 실질적으로 상당하는 제약을 두지 않는 용어로서 사용된다. 단수 형태는 문맥이 달리 명확하기 기술하지 않는 한 대응하는 복수의 지시대상을 포함한다.

[0287] **참고문헌**

Adamopoulou, E., Tenzer, S., Hillen, N., Klug, P., Rota, I.A., Tietz, S., Gebhardt, M., Stevanovic, S., Schild, H., Tolosa, E., et al. (2013). Exploring the MHC-peptide matrix of central tolerance in the human thymus. Nat Commun 4, 2039.

Andreatta, M., and Nielsen, M. (2016). Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics 32, 511-517.

Audemard, E.O., Gendron, P., Feghaly, A., Lavallée, V.P., Hébert, J., Sauvageau, G., and Lemieux, S. (2019). Targeted variant detection using unaligned RNA-Seq reads. Life Sci Alliance 2.

Audoux, J., Philippe, N., Chikhi, R., Salson, M., Gallopin, M., Gabriel, M., Le Coz, J., Drouineau, E., Commes, T., and Gautheret, D. (2017). DE-kupl: exhaustive capture of biological variation in RNA-seq data through k-mer decomposition. Genome Biology 18, 243.

Avigan, D., and Rosenblatt, J. (2018). Vaccine therapy in hematologic malignancies. Blood 131, 2640-2650.

Bamezai, S., Rawat, V.P., and Buske, C. (2012). Concise review: The Piwi-piRNA axis: pivotal beyond transposon silencing. Stem Cells 30, 2603-2611.

Berlin, C., Kowalewski, D.J., Schuster, H., Mirza, N., Walz, S., Handel, M., Schmid-Horch, B., Salih, H.R., Kanz, L., Rammensee, H.G., et al. (2015). Mapping the HLA ligandome landscape of acute myeloid leukemia: a targeted approach toward peptide-based immunotherapy. Leukemia 29, 647-659.

Boegel, S., Lower, M., Bukur, T., Sorn, P., Castle, J.C., and Sahin, U. (2018). HLA and proteasome expression body map. BMC Med Genomics 11, 36.

Bouwmeester, R., Gabriels, R., Hulstaert, N., Martens, L., and Degroeve, S. (2020). DeepLC can predict retention times for peptides that carry as-yet unseen modifications. bioRxiv, 2020.2003.2028.013003.

Bray, N.L., Pimentel, H., Melsted, P., and Pachter, L. (2016). Near-optimal probabilistic RNA-seq quantification. Nature Biotechnology 34, 525-527.

Chapuis, A.G., Egan, D.N., Bar, M., Schmitt, T.M., McAfee, M.S., Paulson, K.G., Voillet, V., Gottardo, R., Ragnarsson, G.B., Bleakley, M., et al. (2019). T cell receptor gene therapy targeting WT1 prevents acute myeloid leukemia relapse post-transplant. Nat Med 25, 1064-1072.

Charni, S., de Bettignies, G., Rathore, M.G., Aguiló, J.I., van den Elsen, P.J., Haouzi, D., Hipskind, R.A., Enriquez, J.A., Sanchez-Beato, M., Pardo, J., et al. (2010). Oxidative Phosphorylation Induces De Novo Expression of the MHC Class I in Tumor Cells through the ERK5 Pathway. The Journal of Immunology 185, 3498-3503.

[0288]

Chen, J., Brunner, A.D., Cogan, J.Z., Nunez, J.K., Fields, A.P., Adamson, B., Itzhak, D.N., Li, J.Y., Mann, M., Leonetti, M.D., and Weissman, J.S. (2020). Pervasive functional translation of noncanonical human open reading frames. Science 367, 1140-1146.

Chong, C., Muller, M., Pak, H., Harnett, D., Huber, F., Grun, D., Leleu, M., Auger, A., Arnaud, M., Stevenson, B.J., et al. (2020). Integrated proteogenomic deep sequencing and analytics accurately identify non-canonical peptides in tumor immunopeptidomes. Nat Commun 11, 1293.

Corces, M.R., Buenrostro, J.D., Wu, B., Greenside, P.G., Chan, S.M., Koenig, J.L., Snyder, M.P., Pritchard, J.K., Kundaje, A., Greenleaf, W.J., et al. (2016). Lineage-specific and single-cell chromatin accessibility charts human hematopoiesis and leukemia evolution. Nat Genet 48, 1193-1203.

Cosma, G.L., and Eisenlohr, L.C. (2019). Impact of epitope density on CD8(+) T cell development and function. Mol Immunol 113, 120-125.

Coulie, P.G., Van den Eynde, B.J., van der Bruggen, P., and Boon, T. (2014). Tumour antigens recognized by T lymphocytes: at the core of cancer immunotherapy. Nat Rev Cancer 14, 135-146.

Courcelles, M., Durette, C., Daouda, T., Laverdure, J.P., Vincent, K., Lemieux, S., Perreault, C., and Thibault, P. (2020). MAPDP: A Cloud-Based Computational Platform for Immunopeptidomics Analyses. J Proteome Res.

Danilova, L., Anagnostou, V., Caushi, J.X., Sidhom, J.W., Guo, H., Chan, H.Y., Suri, P., Tam, A., Zhang, J., Asmar, M.E., et al. (2018). The Mutation-Associated Neoantigen Functional Expansion of Specific T Cells (MANAFEST) Assay: A Sensitive Platform for Monitoring Antitumor Immunity. Cancer Immunol Res 6, 888-899.

Daouda, T., Perreault, C., and Lemieux, S. (2016). pyGeno: A Python package for precision medicine and proteogenomics. F1000Res 5, 381.

Di Stasi, A., Jimenez, A.M., Minagawa, K., Al-Obaidi, M., and Rezvani, K. (2015). Review of the Results of WT1 Peptide Vaccination Strategies for Myelodysplastic Syndromes and Acute Myeloid Leukemia from Nine Different Studies. Frontiers in Immunology 6.

Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. Bioinformatics 29, 15-21.

Dutoit, V., Rubio-Godoy, V., Pittet, M.J., Zippelius, A., Dietrich, P.Y., Legal, F.A., Guillaume, P., Romero, P., Cerottini, J.C., Houghten, R.A., et al. (2002). Degeneracy of antigen recognition as the molecular basis for the high frequency of naive A2/Melan-a peptide multimer(+) CD8(+) T cells in humans. J Exp Med 196, 207-216.

Dvinge, H., and Bradley, R.K. (2015). Widespread intron retention diversifies most cancer transcriptomes. Genome medicine 7, 45-45.

[0289]

Efremova, M., Finotello, F., Rieder, D., and Trajanoski, Z. (2017). Neoantigens Generated by Individual Mutations and Their Role in Cancer Immunity and Immunotherapy. Front Immunol 8, 1679.

Egen, J.G., Ouyang, W., and Wu, L.C. (2020). Human Anti-tumor Immunity: Insights from Immunotherapy Clinical Trials. Immunity 52, 36-54.

Ehx, G., and Perreault, C. (2019). Discovery and characterization of actionable tumor antigens. Genome Medicine 11, 29.

Elias, J.E., and Gygi, S.P. (2010). Target-decoy search strategy for mass spectrometry-based proteomics. Methods Mol Biol 604, 55-71.

Eng, J.K., Hoopmann, M.R., Jahan, T.A., Egertson, J.D., Noble, W.S., and MacCoss, M.J. (2015). A deeper look into Comet--implementation and features. J Am Soc Mass Spectrom 26, 1865-1874.

Eppert, K., Takenaka, K., Lechman, E.R., Waldron, L., Nilsson, B., van Galen, P., Metzeler, K.H., Poeppl, A., Ling, V., Beyene, J., et al. (2011). Stem cell gene expression programs influence clinical outcome in human leukemia. Nat Med 17, 1086-1093.

Fennell, K.A., Bell, C.C., and Dawson, M.A. (2019). Epigenetic therapies in acute myeloid leukemia: where to from here? Blood 134, 1891-1901.

Fergusson, J.R., Morgan, M.D., Bruchard, M., Huitema, L., Heesters, B.A., van Unen, V., van Hamburg, J.P., van der Wel, N.N., Picavet, D., Koning, F., et al. (2018). Maturing Human CD127+ CCR7+ PDL1+ Dendritic Cells Express AIRE in the Absence of Tissue Restricted Antigens. Front Immunol 9, 2902.

Figueroa, M.E., Abdel-Wahab, O., Lu, C., Ward, P.S., Patel, J., Shih, A., Li, Y., Bhagwat, N., Vasanthakumar, A., Fernandez, H.F., et al. (2010a). Leukemic IDH1 and IDH2 mutations result in a hypermethylation phenotype, disrupt TET2 function, and impair hematopoietic differentiation. Cancer Cell 18, 553-567.

Figueroa, M.E., Lugthart, S., Li, Y., Erpelinck-Verschueren, C., Deng, X., Christos, P.J., Schifano, E., Booth, J., van Putten, W., Skrabanek, L., et al. (2010b). DNA methylation signatures identify biologically distinct subtypes in acute myeloid leukemia. Cancer Cell 17, 13-27.

Garrison, E., and Marth, G. (2012). Haplotype-based variant detection from short-read sequencing. arXiv: Genomics.

Gaujoux, R., and Seoighe, C. (2010). A flexible R package for nonnegative matrix factorization. BMC Bioinformatics 11, 367.

Greiner, J., Hofmann, S., Schmitt, M., Götz, M., Wiesneth, M., Schrezenmeier, H., Bunjes, D., Döhner, H., and Bullinger, L. (2017). Acute myeloid leukemia with mutated nucleophosmin 1: an immunogenic acute myeloid leukemia subtype and potential candidate for immune checkpoint inhibition. Haematologica 102, e499-e501.

[0290]

Gu, M., Zwiebel, M., Ong, S.H., Boughton, N., Nomdedeu, J., Basheer, F., Nannya, Y., Quiros, P.M., Ogawa, S., Cazzola, M., et al. (2020). RNAmut: robust identification of somatic mutations in acute myeloid leukemia using RNA-sequencing. Haematologica 105, e290-e293.

Gutierrez, S.E., and Romero-Oliva, F.A. (2013). Epigenetic changes: a common theme in acute myelogenous leukemogenesis. J Hematol Oncol 6, 57.

Hardy, M.-P., Vincent, K., and Perreault, C. (2019). The Genomic Landscape of Antigenic Targets for T Cell-Based Leukemia Immunotherapy. Frontiers in Immunology 10, 2934.

He, H., Kondo, Y., Ishiyama, K., Alatrash, G., Lu, S., Cox, K., Qiao, N., Clise-Dwyer, K., St John, L., Sukhumalchandra, P., et al. (2020). Two unique HLA-A*0201 restricted peptides derived from cyclin E as immunotherapeutic targets in leukemia. Leukemia.

Hesnard, L., Legoux, F., Gautreau, L., Moyon, M., Baron, O., Devilder, M.C., Bonneville, M., and Saulquin, X. (2016). Role of the MHC restriction during maturation of antigen-specific human T cells in the thymus. Eur J Immunol 46, 560-569.

Janelle, V., Carli, C., Taillefer, J., Orio, J., and Delisle, J.S. (2015). Defining novel parameters for the optimal priming and expansion of minor histocompatibility antigen-specific T cells in culture. J Transl Med 13, 123.

Jung, N., Dai, B., Gentles, A.J., Majeti, R., and Feinberg, A.P. (2015). An LSC epigenetic signature is largely mutation independent and implicates the HOXA cluster in AML pathogenesis. Nature Communications 6, 8489.

Knaus, H.A., Berglund, S., Hackl, H., Blackford, A.L., Zeidner, J.F., Montiel-Esparza, R., Mukhopadhyay, R., Vanura, K., Blazar, B.R., Karp, J.E., et al. (2018). Signatures of CD8+ T cell dysfunction in AML patients and their reversibility with response to chemotherapy. JCI Insight 3.

Krokhin, O.V. (2006). Sequence-specific retention calculator. Algorithm for peptide retention prediction in ion-pair RP-HPLC: application to 300- and 100-A pore size C18 sorbents. Anal Chem 78, 7785-7795.

Lamoliatte, F., McManus, F.P., Maarifi, G., Chelbi-Alix, M.K., and Thibault, P. (2017). Uncovering the SUMOylation and ubiquitylation crosstalk in human cells using sequential peptide immunopurification. Nat Commun 8, 14109.

Larouche, J.D., Trofimov, A., Hesnard, L., Ehx, G., Zhao, Q., Vincent, K., Durette, C., Gendron, P., Laverdure, J.P., Bonneil, E., et al. (2020). Widespread and tissue-specific expression of endogenous retroelements in human somatic tissues. Genome Med 12, 40.

Laumont, C.M., Daouda, T., Laverdure, J.P., Bonneil, E., Caron-Lizotte, O., Hardy, M.P., Granados, D.P., Durette, C., Lemieux, S., Thibault, P., and Perreault, C. (2016). Global proteogenomic analysis

[0291]

of human MHC class I-associated peptides derived from non-canonical reading frames. Nat Commun 7, 10238.

Laumont, C.M., Vincent, K., Hesnard, L., Audemard, E., Bonneil, E., Laverdure, J.P., Gendron, P., Courcelles, M., Hardy, M.P., Cote, C., et al. (2018). Noncoding regions are the main source of targetable tumor-specific antigens. Sci Transl Med 10.

Lavallee, V.P., Baccelli, I., Krosl, J., Wilhelm, B., Barabe, F., Gendron, P., Boucher, G., Lemieux, S., Marinier, A., Meloche, S., et al. (2015). The transcriptomic landscape and directed chemical interrogation of MLL-rearranged acute myeloid leukemias. Nat Genet 47, 1030-1037.

Lavallee, V.P., Krosl, J., Lemieux, S., Boucher, G., Gendron, P., Pabst, C., Boivin, I., Marinier, A., Guidos, C.J., Meloche, S., et al. (2016). Chemo-genomic interrogation of CEBPA mutated AML reveals recurrent CSF3R mutations and subgroup sensitivity to JAK inhibitors. Blood 127, 3054-3061.

Lawrence, M.S., Stojanov, P., Polak, P., Kryukov, G.V., Cibulskis, K., Sivachenko, A., Carter, S.L., Stewart, C., Mermel, C.H., Roberts, S.A., et al. (2013). Mutational heterogeneity in cancer and the search for new cancer-associated genes. Nature 499, 214-218.

Legat, A., Maby-El Hajjami, H., Baumgaertner, P., Cagnon, L., Abed Maillard, S., Geldhof, C., Iancu, E.M., Lebon, L., Guillaume, P., Dojcinovic, D., et al. (2016). Vaccination with LAG-3Ig (IMP321) and Peptides Induces Specific CD4 and CD8 T-Cell Responses in Metastatic Melanoma Patients--Report of a Phase I/IIa Clinical Trial. Clin Cancer Res 22, 1330-1340.

Lewinsky, H., Barak, A.F., Huber, V., Kramer, M.P., Radomir, L., Sever, L., Orr, I., Mirkin, V., Dezorella, N., Shapiro, M., et al. (2018). CD84 regulates PD-1/PD-L1 expression and function in chronic lymphocytic leukemia. J Clin Invest 128, 5465-5478.

Ley, T.J., Miller, C., Ding, L., Raphael, B.J., Mungall, A.J., Robertson, A., Hoadley, K., Triche, T.J., Jr., Laird, P.W., Baty, J.D., et al. (2013). Genomic and epigenomic landscapes of adult de novo acute myeloid leukemia. N Engl J Med 368, 2059-2074.

Li, B., Li, T., Wang, B., Dou, R., Zhang, J., Liu, J.S., and Liu, X.S. (2017). Ultrasensitive detection of TCR hypervariable-region sequences in solid-tissue RNA-seq data. Nat Genet 49, 482-483.

Li, H. (2011). A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. Bioinformatics 27, 2987-2993.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. Bioinformatics 25, 2078-2079.

Li, S., Garrett-Bakelman, F.E., Chung, S.S., Sanders, M.A., Hricik, T., Rapaport, F., Patel, J., Dillon, R., Vijay, P., Brown, A.L., et al. (2016). Distinct evolution and dynamics of epigenetic and genetic heterogeneity in acute myeloid leukemia. Nat Med 22, 792-799.

[0292]

Loffler, M.W., Mohr, C., Bichmann, L., Freudenmann, L.K., Walzer, M., Schroeder, C.M., Trautwein, N., Hilke, F.J., Zinser, R.S., Muhlenbruch, L., et al. (2019). Multi-omics discovery of exome-derived neoantigens in hepatocellular carcinoma. Genome Med 11, 28.

Logtenberg, M.E.W., Scheeren, F.A., and Schumacher, T.N. (2020). The CD47-SIRPα Immune Checkpoint. Immunity 52, 742-752.

Luo, K., Yuan, J., Shan, Y., Li, J., Xu, M., Cui, Y., Tang, W., Wan, B., Zhang, N., Wu, Y., and Yu, L. (2006). Activation of transcriptional activities of AP1 and SRE by a novel zinc finger protein ZNF445. Gene 367, 89-100.

Macrae, T., Sargeant, T., Lemieux, S., Hebert, J., Deneault, E., and Sauvageau, G. (2013). RNA-Seq reveals spliceosome and proteasome genes as most consistent transcripts in human cancer cells. PLoS One 8, e72884.

Maere, S., Heymans, K., and Kuiper, M. (2005). BiNGO: a Cytoscape plugin to assess overrepresentation of gene ontology categories in biological networks. Bioinformatics 21, 3448-3449.

Maiga, A., Lemieux, S., Pabst, C., Lavallee, V.P., Bouvier, M., Sauvageau, G., and Hebert, J. (2016). Transcriptome analysis of G protein-coupled receptors in distinct genetic subgroups of acute myeloid leukemia: identification of potential disease-specific targets. Blood Cancer J 6, e431.

Marcais, G., and Kingsford, C. (2011). A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. Bioinformatics 27, 764-770.

Maslak, P.G., Dao, T., Bernal, Y., Chanel, S.M., Zhang, R., Frattini, M., Rosenblat, T., Jurcic, J.G., Brentjens, R.J., Arcila, M.E., et al. (2018). Phase 2 trial of a multivalent WT1 peptide vaccine (galinpepimut-S) in acute myeloid leukemia. Blood Adv 2, 224-234.

Merico, D., Isserlin, R., Stueker, O., Emili, A., and Bader, G.D. (2010). Enrichment map: a network-based method for gene-set enrichment visualization and interpretation. PLoS One 5, e13984.

Middleton, R., Gao, D., Thomas, A., Singh, B., Au, A., Wong, J.J., Bomane, A., Cosson, B., Eyras, E., Rasko, J.E., and Ritchie, W. (2017). IRFinder: assessing the impact of intron retention on mammalian gene expression. Genome Biol 18, 51.

Ntziachristos, P., Abdel-Wahab, O., and Aifantis, I. (2016). Emerging concepts of epigenetic dysregulation in hematological malignancies. Nat Immunol 17, 1016-1024.

Ogishi, M., and Yotsuyanagi, H. (2019). Quantitative Prediction of the Landscape of T Cell Epitope Immunogenicity in Sequence Space. Frontiers in Immunology 10.

Pabst, C., Bergeron, A., Lavallee, V.P., Yeh, J., Gendron, P., Norddahl, G.L., Krosl, J., Boivin, I., Deneault, E., Simard, J., et al. (2016). GPR56 identifies primary human acute myeloid leukemia cells with high repopulating potential in vivo. Blood 127, 2018-2027.

[0293]

Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V.I., Paschka, P., Roberts, N.D., Potter, N.E., Heuser, M., Thol, F., Bolli, N., et al. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. N Engl J Med 374, 2209-2221.

Pearson, H., Daouda, T., Granados, D.P., Durette, C., Bonneil, E., Courcelles, M., Rodenbrock, A., Laverdure, J.P., Cote, C., Mader, S., et al. (2016). MHC class I-associated peptides derive from selective regions of the human genome. J Clin Invest 126, 4690-4701.

Qazilbash, M.H., Wieder, E., Thall, P.F., Wang, X., Rios, R., Lu, S., Kanodia, S., Ruisaard, K.E., Giralt, S.A., Estey, E.H., et al. (2017). PR1 peptide vaccine induces specific immunity with clinical responses in myeloid malignancies. Leukemia 31, 697-704.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26, 841-842.

Rashidi, A., and Walter, R.B. (2016). Antigen-specific immunotherapy for acute myeloid leukemia: where are we now, and where do we go from here? Expert Rev Hematol 9, 335-350.

Ritchie, M.E., Phipson, B., Wu, D., Hu, Y., Law, C.W., Shi, W., and Smyth, G.K. (2015). limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 43, e47.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. Nat Biotechnol 29, 24-26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26, 139-140.

Rogers, M.F., Shihab, H.A., Mort, M., Cooper, D.N., Gaunt, T.R., and Campbell, C. (2018). FATHMM-XF: accurate prediction of pathogenic point mutations via extended features. Bioinformatics 34, 511-513.

Sarkizova, S., Klaeger, S., Le, P.M., Li, L.W., Oliveira, G., Keshishian, H., Hartigan, C.R., Zhang, W., Braun, D.A., Ligon, K.L., et al. (2020). A large peptidome dataset improves HLA class I epitope prediction across most of the human population. Nat Biotechnol 38, 199-209.

Schulz, W.A., Steinhoff, C., and Florl, A.R. (2006). Methylation of endogenous human retroelements in health and disease. Curr Top Microbiol Immunol 310, 211-250.

Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., and Ideker, T. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 13, 2498-2504.

Shao, W., Pedrioli, P.G.A., Wolski, W., Scurtescu, C., Schmid, E., Vizcaino, J.A., Courcelles, M., Schuster, H., Kowalewski, D., Marino, F., et al. (2018). The SysteMHC Atlas project. Nucleic Acids Res 46, D1237-d1247.

[0294]

Shlush, L.I., Mitchell, A., Heisler, L., Abelson, S., Ng, S.W.K., Trotman-Grant, A., Medeiros, J.J.F., Rao-Bhatia, A., Jaciw-Zurakowsky, I., Marke, R., et al. (2017). Tracing the origins of relapse in acute myeloid leukaemia to stem cells. Nature 547, 104-108.

Smart, A.C., Margolis, C.A., Pimentel, H., He, M.X., Miao, D., Adeegbe, D., Fugmann, T., Wong, K.-K., and Van Allen, E.M. (2018). Intron retention is a source of neoepitopes in cancer. Nature Biotechnology 36, 1056-1058.

Smith, C.C., Selitsky, S.R., Chai, S., Armistead, P.M., Vincent, B.G., and Serody, J.S. (2019). Alternative tumour-specific antigens. Nat Rev Cancer 19, 465-478.

Springer, I., Besser, H., Tickotsky-Moskovitz, N., Dvorkin, S., and Louzoun, Y. (2020). Prediction of Specific TCR-Peptide Binding From Large Dictionaries of TCR-Peptide Pairs. Front Immunol 11, 1803.

Szolek, A., Schubert, B., Mohr, C., Sturm, M., Feldhahn, M., and Kohlbacher, O. (2014). OptiType: precision HLA typing from next-generation sequencing data. Bioinformatics 30, 3310-3316.

Takahashi, N., Coluccio, A., Thorball, C.W., Planet, E., Shi, H., Offner, S., Turelli, P., Imbeault, M., Ferguson-Smith, A.C., and Trono, D. (2019). ZNF445 is a primary regulator of genomic imprinting. Genes Dev 33, 49-54.

Tamura, H., Dan, K., Tamada, K., Nakamura, K., Shioi, Y., Hyodo, H., Wang, S.D., Dong, H., Chen, L., and Ogata, K. (2005). Expression of functional B7-H2 and B7.2 costimulatory molecules and their prognostic implications in de novo acute myeloid leukemia. Clin Cancer Res 11, 5708-5717.

Tate, J.G., Bamford, S., Jubb, H.C., Sondka, Z., Beare, D.M., Bindal, N., Boutselakis, H., Cole, C.G., Creatore, C., Dawson, E., et al. (2018). COSMIC: the Catalogue Of Somatic Mutations In Cancer. Nucleic Acids Research 47, D941-D947.

Thomsen, M., Lundegaard, C., Buus, S., Lund, O., and Nielsen, M. (2013). MHCcluster, a method for functional clustering of MHC molecules. Immunogenetics 65, 655-665.

Toffalori, C., Zito, L., Gambacorta, V., Riba, M., Oliveira, G., Bucci, G., Barcella, M., Spinelli, O., Greco, R., Crucitti, L., et al. (2019). Immune signature drives leukemia escape and relapse after hematopoietic cell transplantation. Nat Med 25, 603-611.

Tripathi, S.C., Peters, H.L., Taguchi, A., Katayama, H., Wang, H., Momin, A., Jolly, M.K., Celiktas, M., Rodriguez-Canales, J., Liu, H., et al. (2016). Immunoproteasome deficiency is a feature of non-small cell lung cancer with a mesenchymal phenotype and is associated with a poor outcome. Proc Natl Acad Sci U S A 113, E1555-1564.

van der Lee, D.I., Reijmers, R.M., Honders, M.W., Hagedoorn, R.S., de Jong, R.C., Kester, M.G., van der Steen, D.M., de Ru, A.H., Kweekel, C., Bijen, H.M., et al. (2019). Mutated nucleophosmin 1 as immunotherapy target in acute myeloid leukemia. J Clin Invest 129, 774-785.

[0295]

Vasu, S., Kohlschmidt, J., Mrozek, K., Eisfeld, A.K., Nicolet, D., Sterling, L.J., Becker, H., Metzeler, K.H., Papaioannou, D., Powell, B.L., et al. (2018). Ten-year outcome of patients with acute myeloid leukemia not treated with allogeneic transplantation in first complete remission. Blood Adv 2, 1645-1650.

Vizcaino, J.A., Csordas, A., del-Toro, N., Dianes, J.A., Griss, J., Lavidas, I., Mayer, G., Perez-Riverol, Y., Reisinger, F., Ternent, T., et al. (2016). 2016 update of the PRIDE database and its related tools. Nucleic Acids Res 44, D447-456.

Wang, E., Lu, S.X., Pastore, A., Chen, X., Imig, J., Chun-Wei Lee, S., Hockemeyer, K., Ghebrechristos, Y.E., Yoshimi, A., Inoue, D., et al. (2019a). Targeting an RNA-Binding Protein Network in Acute Myeloid Leukemia. Cancer Cell 35, 369-384.e367.

Wang, Q., Ding, H., He, Y., Li, X., Cheng, Y., Xu, Q., Yang, Y., Liao, G., Meng, X., Huang, C., and Li, J. (2019b). NLRC5 mediates cell proliferation, migration, and invasion by regulating the Wnt/beta-catenin signalling pathway in clear cell renal cell carcinoma. Cancer Lett 444, 9-19.

Whiteway, A., Corbett, T., Anderson, R., Macdonald, I., and Prentice, H.G. (2003). Expression of co-stimulatory molecules on acute myeloid leukaemia blasts may effect duration of first remission. Br J Haematol 120, 442-451.

Wong, J.J.L., Gao, D., Nguyen, T.V., Kwok, C.-T., van Geldermalsen, M., Middleton, R., Pinello, N., Thoeng, A., Nagarajah, R., Holst, J., et al. (2017). Intron retention is regulated by altered MeCP2-mediated splicing factor recruitment. Nature Communications 8, 15134.

Wu, T.D., and Nacu, S. (2010). Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics 26, 873-881.

Yang, L., Rau, R., and Goodell, M.A. (2015). DNMT3A in haematological malignancies. Nat Rev Cancer 15, 152-165.

Zhang, J., Hu, X., Wang, J., Sahu, A.D., Cohen, D., Song, L., Ouyang, Z., Fan, J., Wang, B., Fu, J., et al. (2019). Immune receptor repertoires in pediatric and adult acute myeloid leukemia. Genome Med 11, 73.

Zhao, Q., Laverdure, J.P., Lanoix, J., Durette, C., Cote, C., Bonneil, E., Laumont, C.M., Gendron, P., Vincent, K., Courcelles, M., et al. (2020). Proteogenomics Uncovers a Vast Repertoire of Shared Tumor-Specific Antigens in Ovarian Cancer. Cancer Immunol Res.

Zhi, H., Ning, S., Li, X., Li, Y., Wu, W., and Li, X. (2014). A novel reannotation strategy for dissecting DNA methylation patterns of human long intergenic non-coding RNAs in cancers. Nucleic Acids Res 42, 8258-8270.

Zhou, J., and Chng, W.J. (2017). Aberrant RNA splicing and mutations in spliceosome complex in acute myeloid leukemia. Stem Cell Investig 4, 6.

[0296]

Zhou, Y., Lu, Y., and Tian, W. (2012). Epigenetic features are significantly associated with alternative splicing. BMC Genomics 13, 123.

Wilson CS, Davidson GS, Martin SB, Andries E, Potter J, Harvey R, et al. Gene expression profiling of adult acute myeloid leukemia identifies novel biologic clusters for risk classification and outcome prediction. Blood 2006; 108(2): 685-696.

Liu L, Xu F, Chang C-K, He Q, Wu L-Y, Zhang Z, et al. MYCN contributes to the malignant characteristics of erythroleukemia through EZH2-mediated epigenetic repression of p21. Cell Death & Disease 2017 2017/10/01; 8(10): e3126-e3126.

Yang X, Lu B, Sun X, Han C, Fu C, Xu K, et al. ANP32A regulates histone H3 acetylation and promotes leukemogenesis. Leukemia 2018 2018/07/01; 32(7): 1587-1597.

de Sá Machado Araújo G, da Silva Francisco Junior R, dos Santos Ferreira C, Mozer Rodrigues PT, Terra Machado D, Louvain de Souza T, et al. Maternal 5mCpG Imprints at the PARD6G-AS1 and GCSAML Differentially Methylated Regions Are Decoupled From Parent-of-Origin Expression Effects in Multiple Human Tissues. Frontiers in Genetics 2018 2018-March-01; 9(36).

Lagus H, Klaas M, Juteau S, Elomaa O, Kere J, Vuola J, et al. Discovery of increased epidermal DNAH10 expression after regeneration of dermis in a randomized with-in person trial — reflections on psoriatic inflammation. Scientific Reports 2019 2019/12/13; 9(1): 19136.

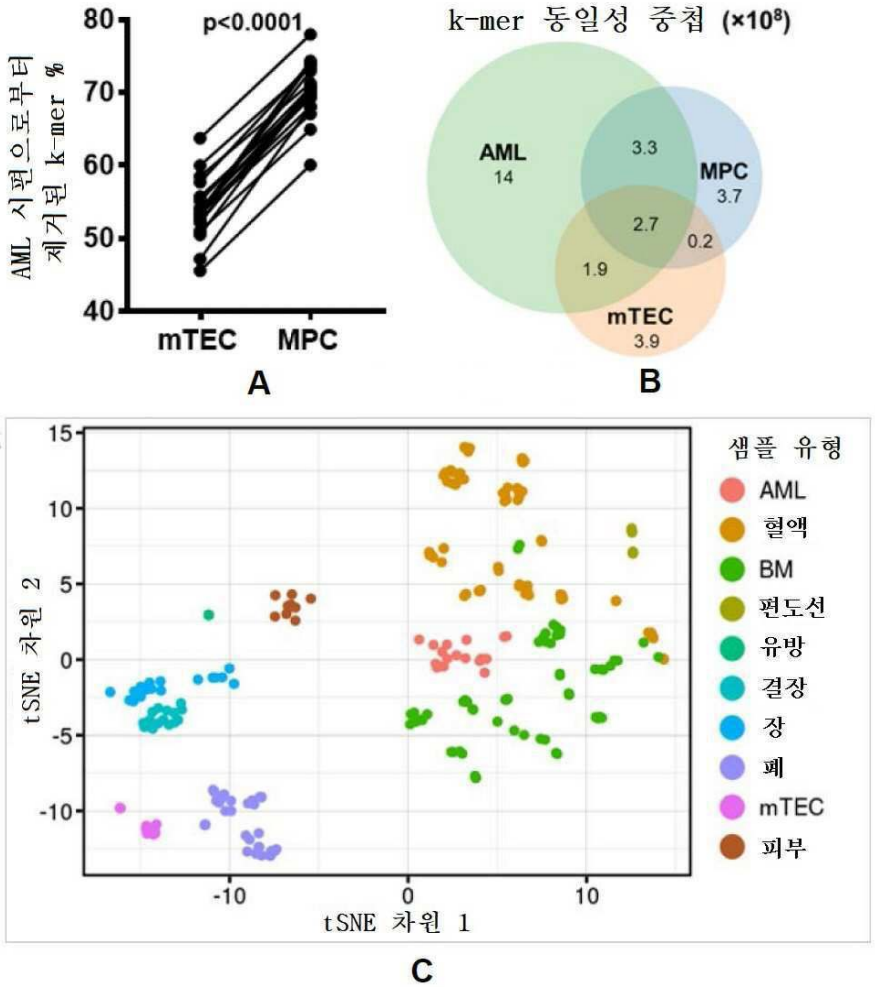
Zhang X, Dong W, Zhou H, Li H, Wang N, Miao X, et al. alpha-2,8-Sialyltransferase Is Involved in the Development of Multidrug Resistance via PI3K/Akt Pathway in Human Chronic Myeloid Leukemia. IUBMB Life 2015 Feb; 67(2): 77-87.

Johnson KD, Kong G, Gao X, Chang Y-I, Hewitt KJ, Sanalkumar R, et al. Cis-regulatory mechanisms governing stem and progenitor cell transitions. Science advances 2015; 1(8): e1500503-e1500503.

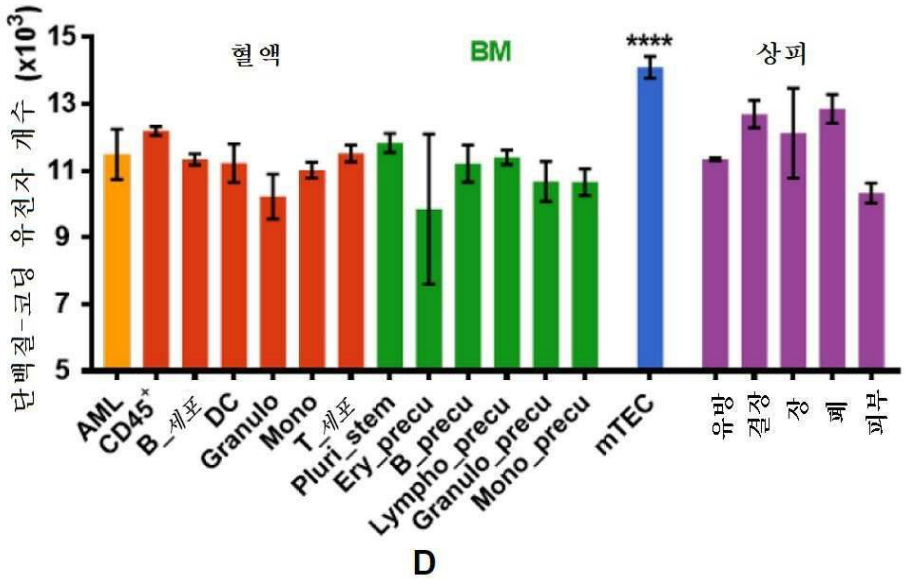
[0297]

도면

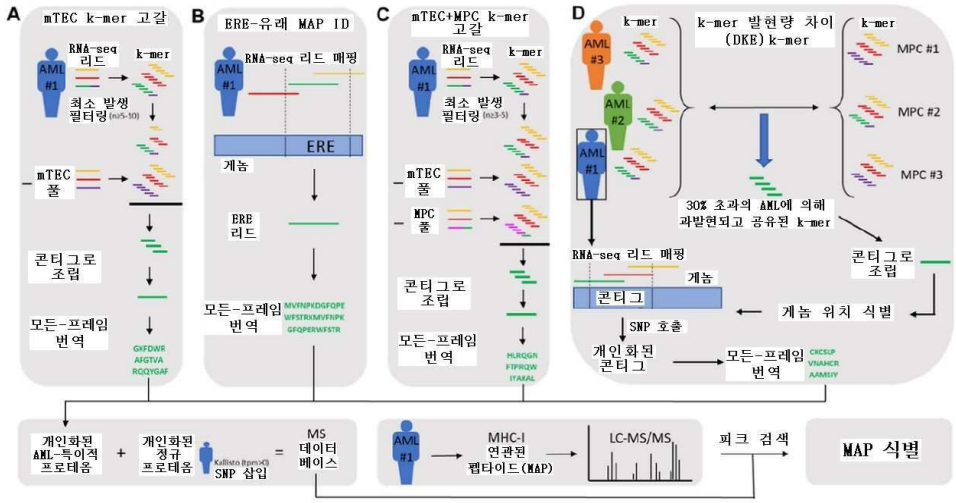
도면1ac



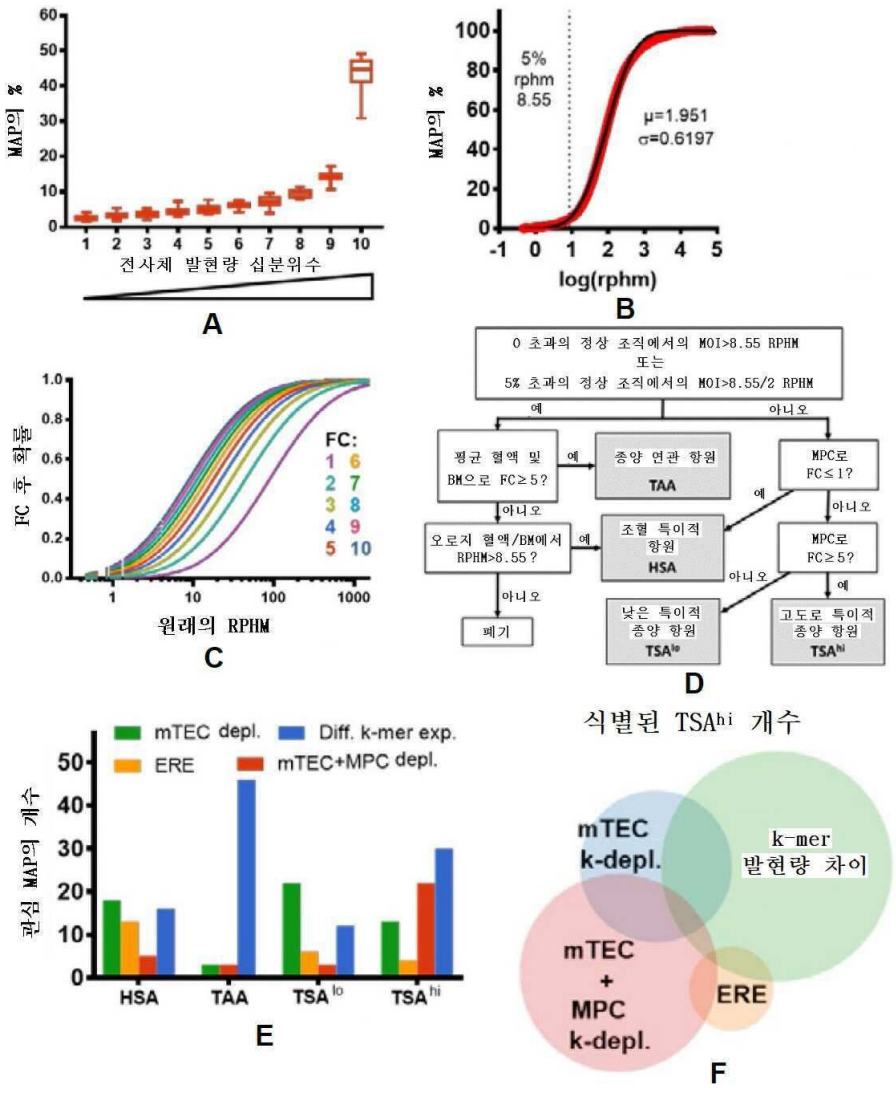
도면1d



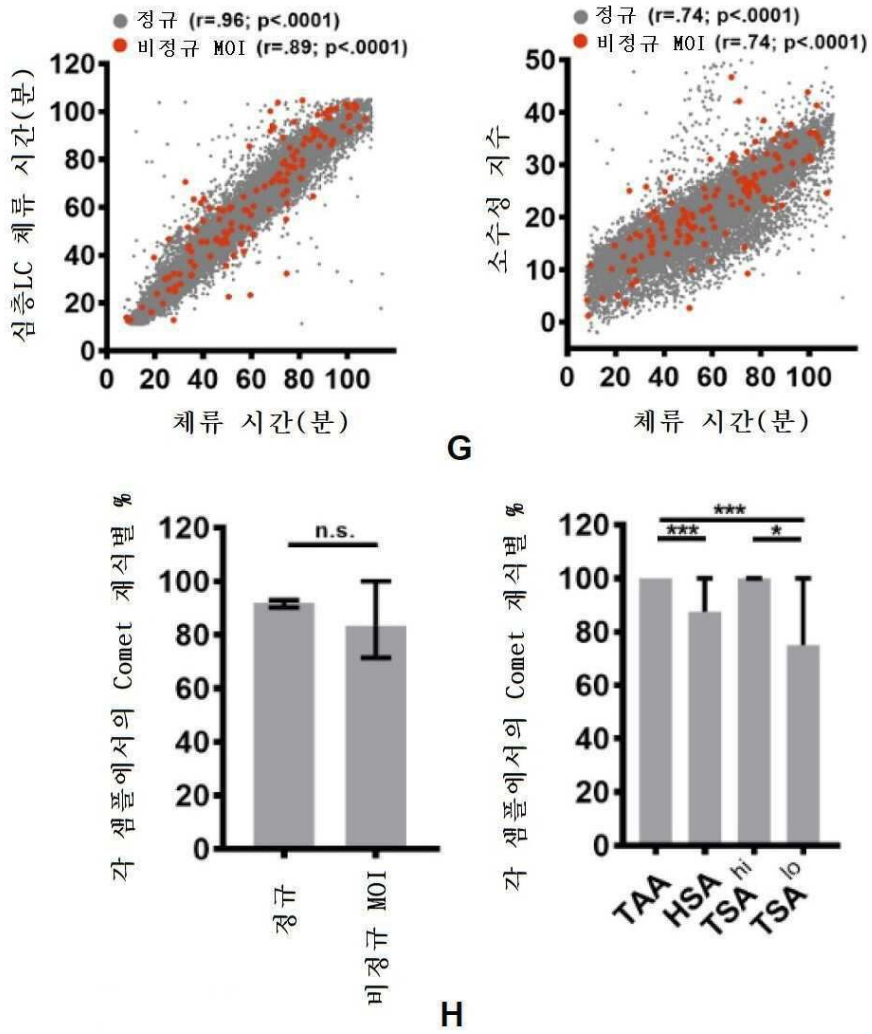
도면2


A
mTEC k-mer 고갈
RNA-seq 리드
k-mer
AML #1
최소 발생 필터링
mTEC 풀
콘티그로 조립
모든-프레임 번역
B
ERE-유래 MAP ID
RNA-seq 리드 매핑
ERE
게놈
ERE 리드
모든-프레임 번역
C
mTEC+MPC k-mer 고갈
mTEC 풀
MPC 풀
콘티그로 조립
모든-프레임 번역
D
k-mer
k-mer 발현량 차이 (DKE) k-mer
AML #3
AML #2
AML #1
MPC #1
MPC #2
MPC #3
30% 초과의 AML에 의해 과발현되고 공유된 k-mer
RNA-seq 리드 매핑
게놈
콘티그
콘티그로 조립
게놈 위치 식별
SNP 호출
개인화된 콘티그
모든-프레임 번역
개인화된 AML-특이적 프로테옴
개인화된 정규 프로테옴
SNP 삽입
MS 데이터 베이스
MHC-I 연관된 펩타이드(MAP)
LC-MS/MS
피크 검색
MAP 식별

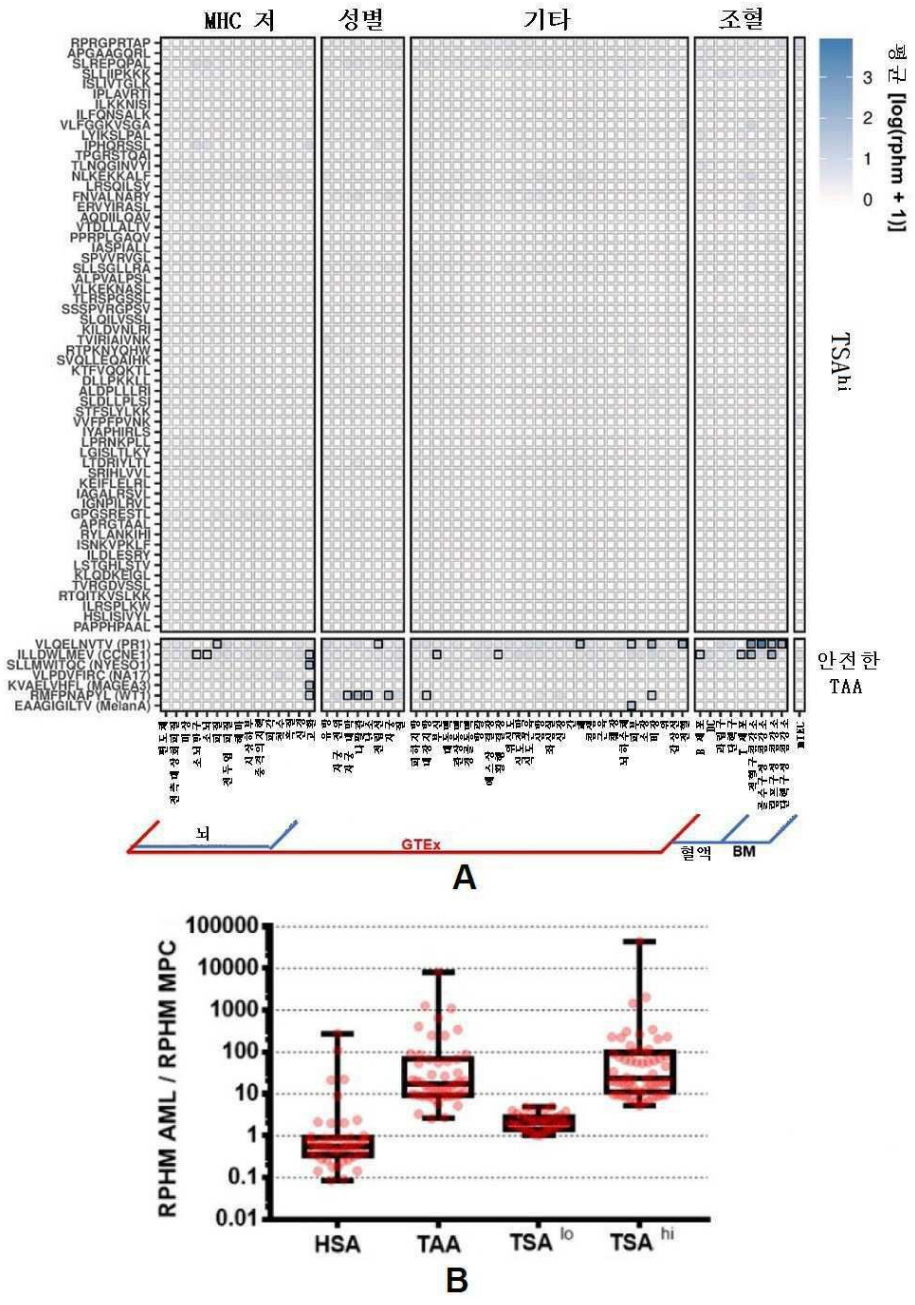
도면3af


MAP의 %
전사체 발현량 십분위수
A
5%
rphm
8.55
μ=1.951
σ=0.6197
log(rphm)
B
FC 후 확률
FC:
원래의 RPHM
C
0 초과의 정상 조직에서의 MOI>8.55 RPHM
또는
5% 초과의 정상 조직에서의 MOI>8.55/2 RPHM
예
아니오
평균 혈액 및 BM으로 FC≥5?
종양 연관 항원
TAA
MPC로 FC≤1?
오로지 혈액/BM에서 RPHM>8.55?
조혈 특이적 항원
HSA
MPC로 FC≥5?
폐기
낮은 특이적 종양 항원
TSAlo
고도로 특이적 종양 항원
TSAhi
D
식별된 TSAhi 개수
mTEC depl.
Diff. k-mer exp.
ERE
mTEC+MPC depl.
관심 MAP의 개수
HSA
TAA
TSAlo
TSAhi
E
mTEC k-depl.
k-mer 발현량 차이
mTEC + MPC k-depl.
ERE
F

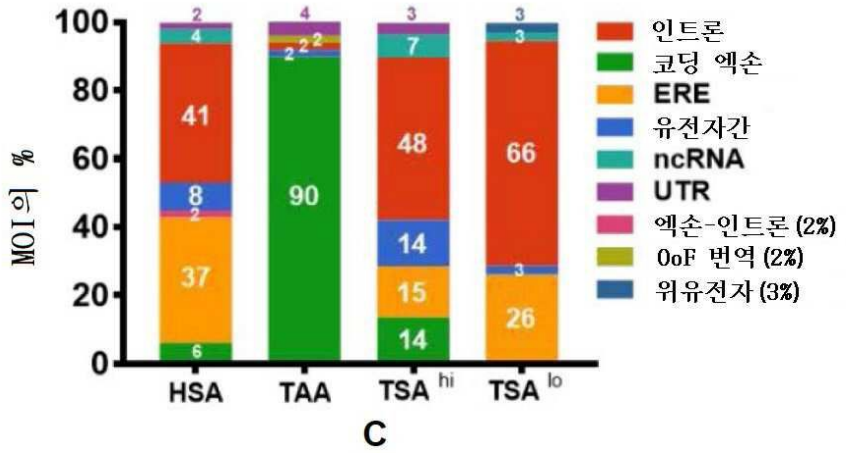
도면3gh


정규 (r=.96; p<.0001)
비정규 MOI (r=.89; p<.0001)
심층LC 체류 시간(분)
체류 시간(분)
정규 (r=.74; p<.0001)
비정규 MOI (r=.74; p<.0001)
소수성 지수
체류 시간(분)
G
각 샘플에서의 Comet 제식별 %
n.s.
정규
비정규 MOI
각 샘플에서의 Comet 제식별 %
***
***
*
TAA
HSA
TSA hi
TSA lo
H

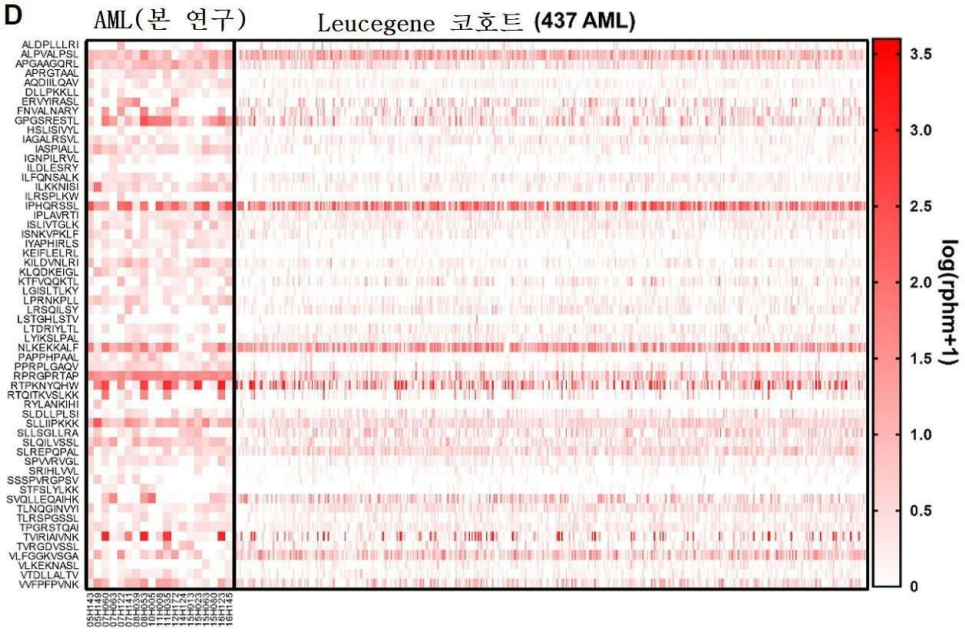
도면4ab


MHC 저
성별
기타
조혈
발현 분율 [log(rphm + 1)]
3
2
1
0
TSA hi
안전한 TAA
VLQELNVTV (PB1)
ILLDWLMEV (CCNE1)
SLLMWITQC (NYESO1)
VLPDVFIRC (NA17)
KVAELVHFL (MAGEA3)
FMFPNAPYL (WT1)
EAAGIGILTV (Melan/A)
뇌
GTEx
혈액
BM
A
RPHM AML / RPHM MPC
100000
10000
1000
100
10
1
0.1
0.01
HSA
TAA
TSA lo
TSA hi
B

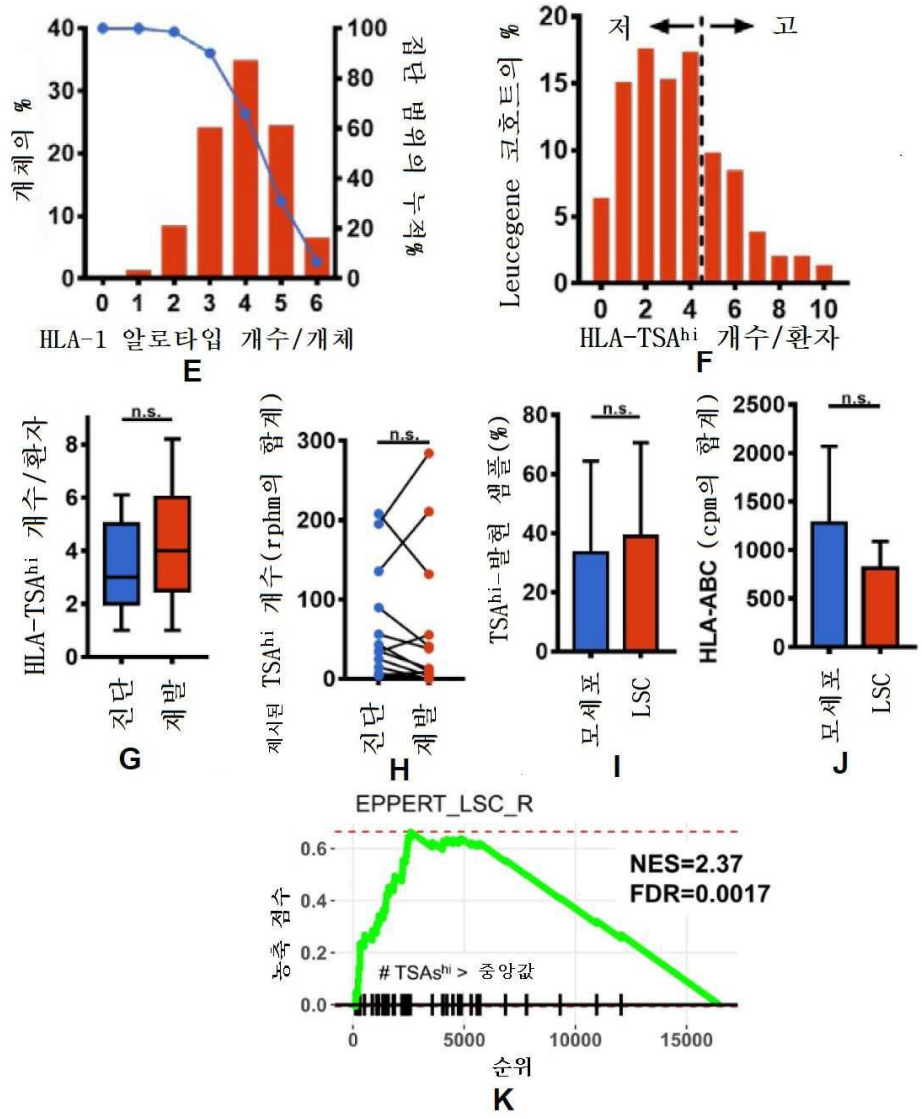
도면4c


100
80
60
40
20
0
MOI의 %
HSA
TAA
TSA hi
TSA lo
C
인트론
코딩 엑손
ERE
유전자간
ncRNA
UTR
엑손-인트론 (2%)
OoF 번역 (2%)
위유전자 (3%)

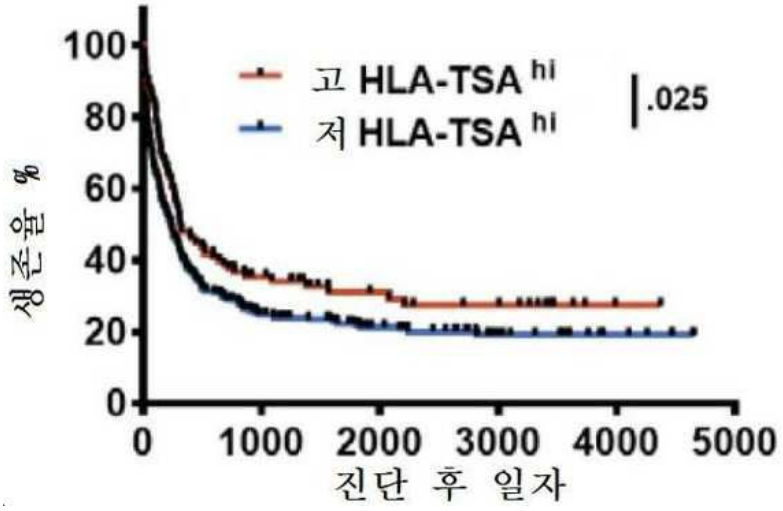
도면4d


D
AML(본 연구)
Leucegene 코호트 (437 AML)
log(rphm+1)
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0

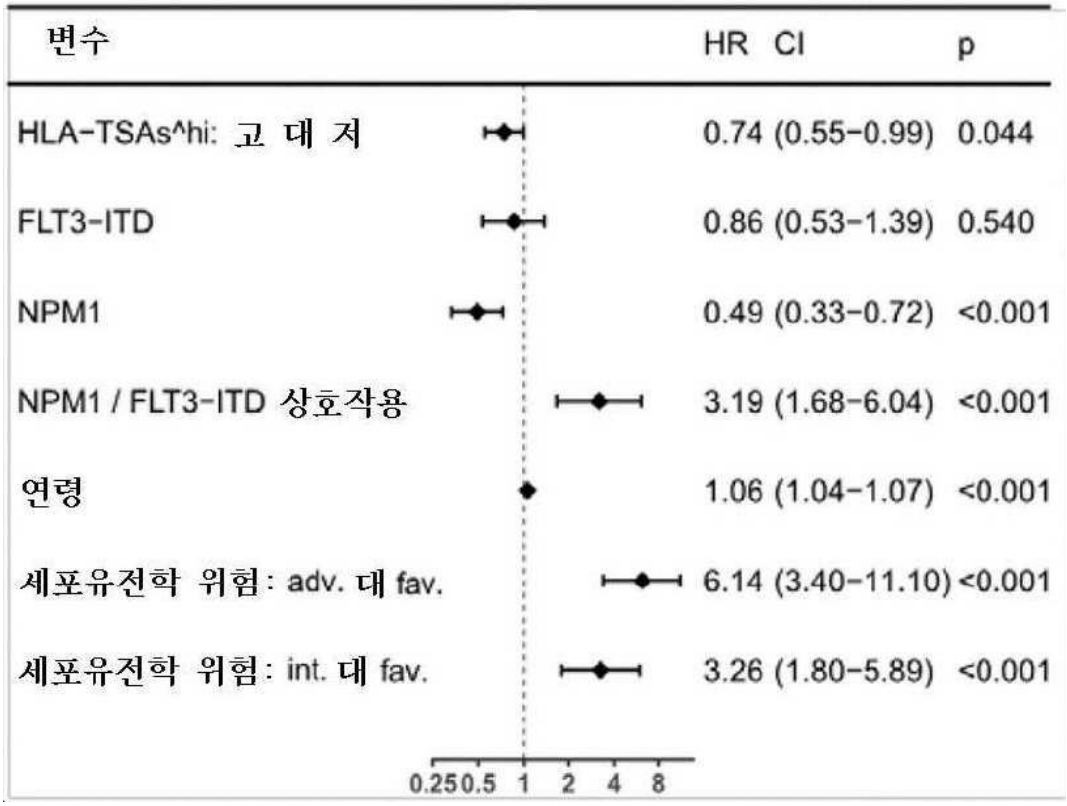
도면4ek


개체의 %
환자의 누적 %
HLA-1 알로타입 개수/개체
E
Leucegene 코호트의 %
저
고
HLA-TSA^hi 개수/환자
F
HLA-TSA^hi 개수/환자
n.s.
진단
재발
G
제시된 TSA^hi 개수(rphm의 합계)
H
TSA^hi-발현 샘플(%)
모세포
LSC
I
HLA-ABC (cpm의 합계)
J
EPPERT_LSC_R
농축 점수
NES=2.37
FDR=0.0017
# TSAs^hi > 중앙값
순위
K

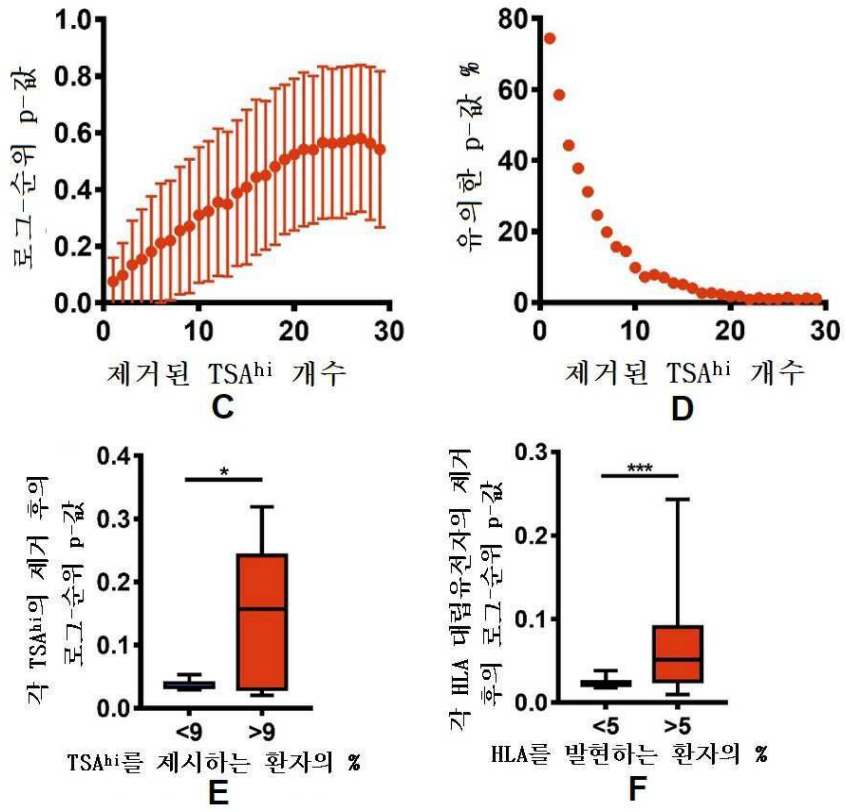
도면5a


생존율 %
고 HLA-TSA hi
저 HLA-TSA hi
.025
진단 후 일자

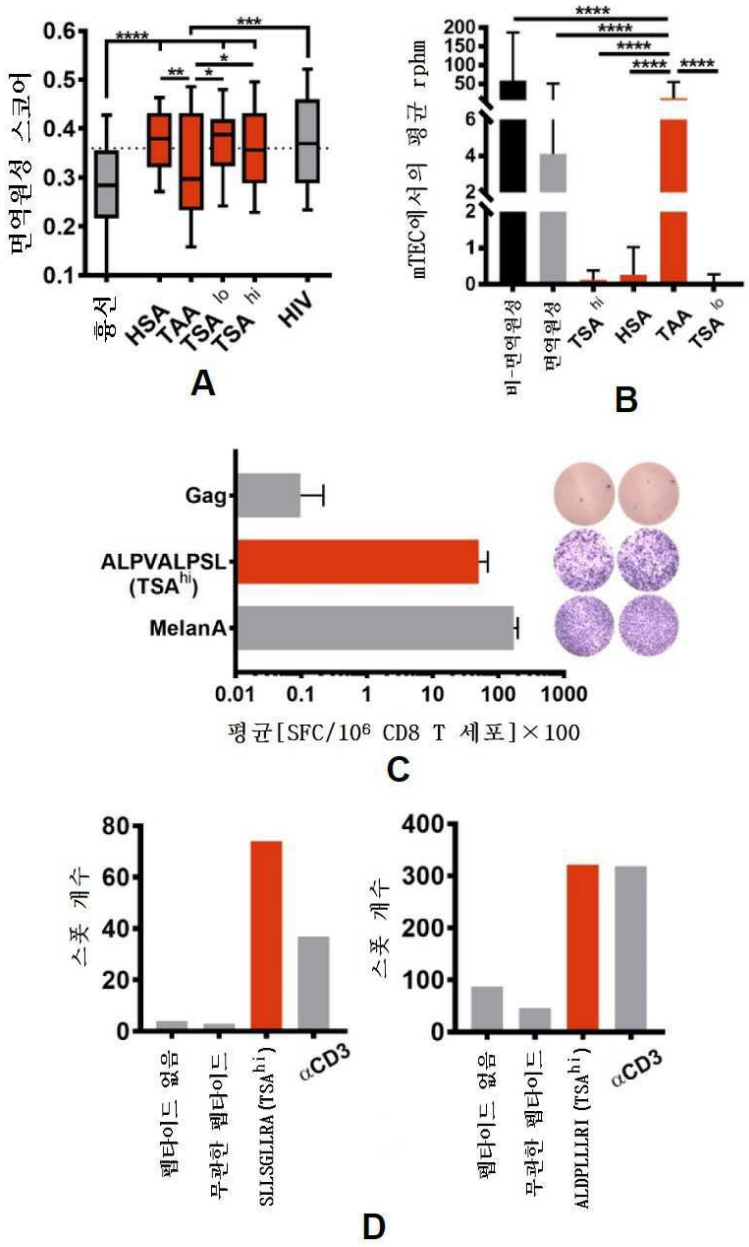
도면5b


변수
HR CI
p
HLA-TSAs^hi: 고 대 저
0.74 (0.55-0.99) 0.044
FLT3-ITD
0.86 (0.53-1.39) 0.540
NPM1
0.49 (0.33-0.72) <0.001
NPM1 / FLT3-ITD 상호작용
3.19 (1.68-6.04) <0.001
연령
1.06 (1.04-1.07) <0.001
세포유전학 위험: adv. 대 fav.
6.14 (3.40-11.10) <0.001
세포유전학 위험: int. 대 fav.
3.26 (1.80-5.89) <0.001
0.25 0.5 1 2 4 8

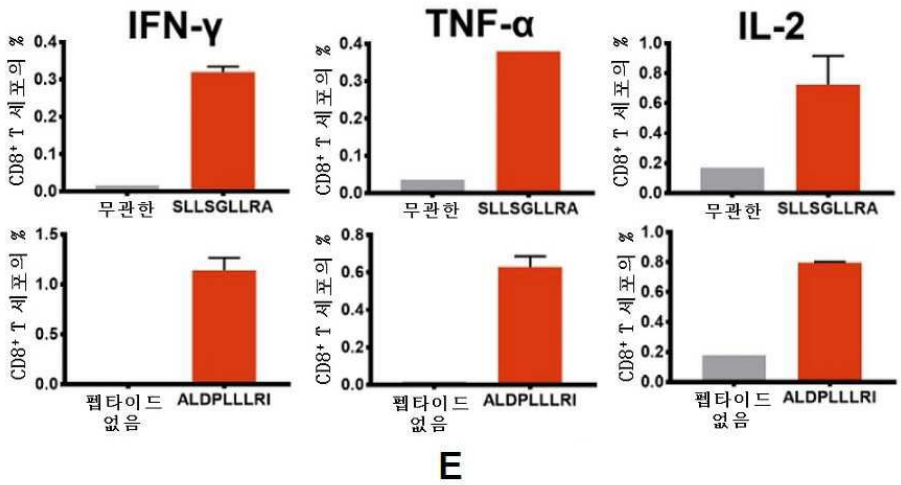
도면5cf

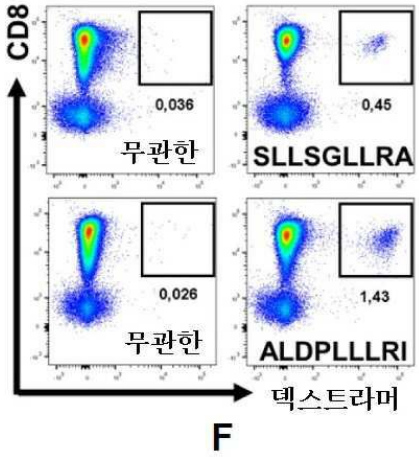

로그-순위 p-값
제거된 TSAhi 개수
C
유의한 p-값 %
제거된 TSAhi 개수
D
각 TSAhi의 제거 후의 로그-순위 p-값
TSAhi를 제시하는 환자의 %
<9
>9
*
E
각 HLA 대립유전자의 제거 후의 로그-순위 p-값
HLA를 발현하는 환자의 %
<5
>5
***
F

도면6ad

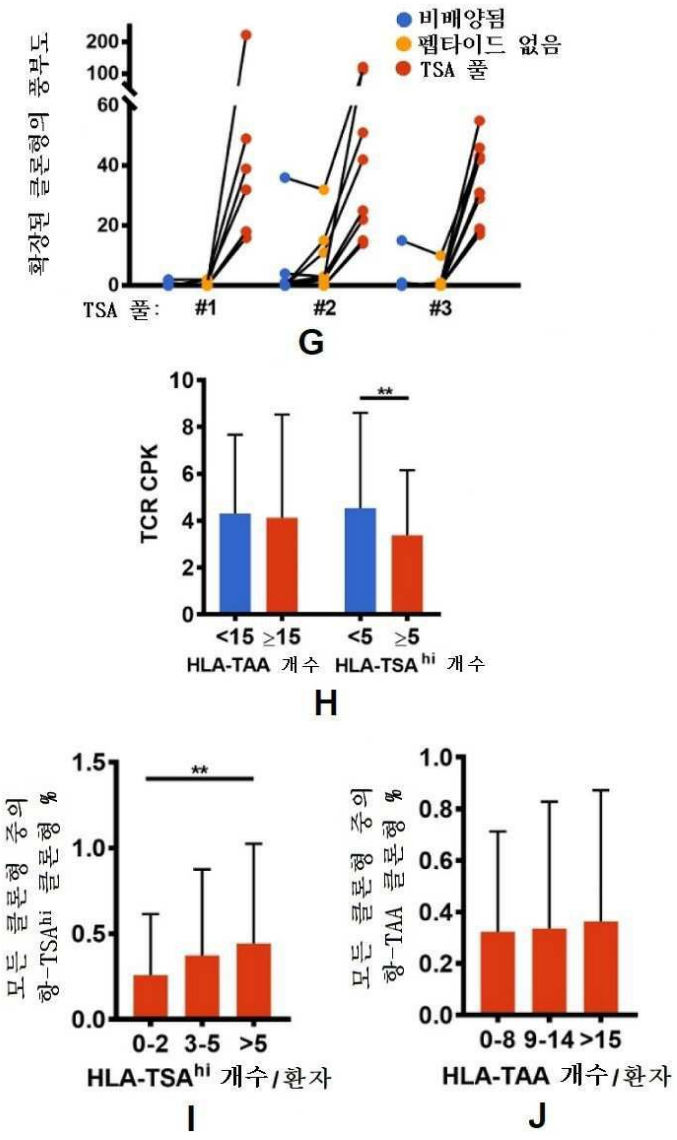

면역원성 스코어
0.6
0.5
0.4
0.3
0.2
0.1
****
***
**
*
*
흉선
HSA
TAA
TSA lo
TSA hi
HIV
A
mTEC에서의 평균 rphm
200
150
100
50
6
4
2
2
1
0
****
****
****
****
****
비-면역원성
면역원성
TSA hi
HSA
TAA
TSA lo
B
Gag
ALPVALPSL (TSA hi)
MelanA
0.01
0.1
1
10
100
1000
평균[SFC/10⁶ CD8 T 세포]×100
C
스폿 개수
80
60
40
20
0
펩타이드 없음
무관한 펩타이드
SLLSGLLRA (TSA hi)
αCD3
스폿 개수
400
300
200
100
0
펩타이드 없음
무관한 펩타이드
ALDPLLLRI (TSA hi)
αCD3
D

도면6ef

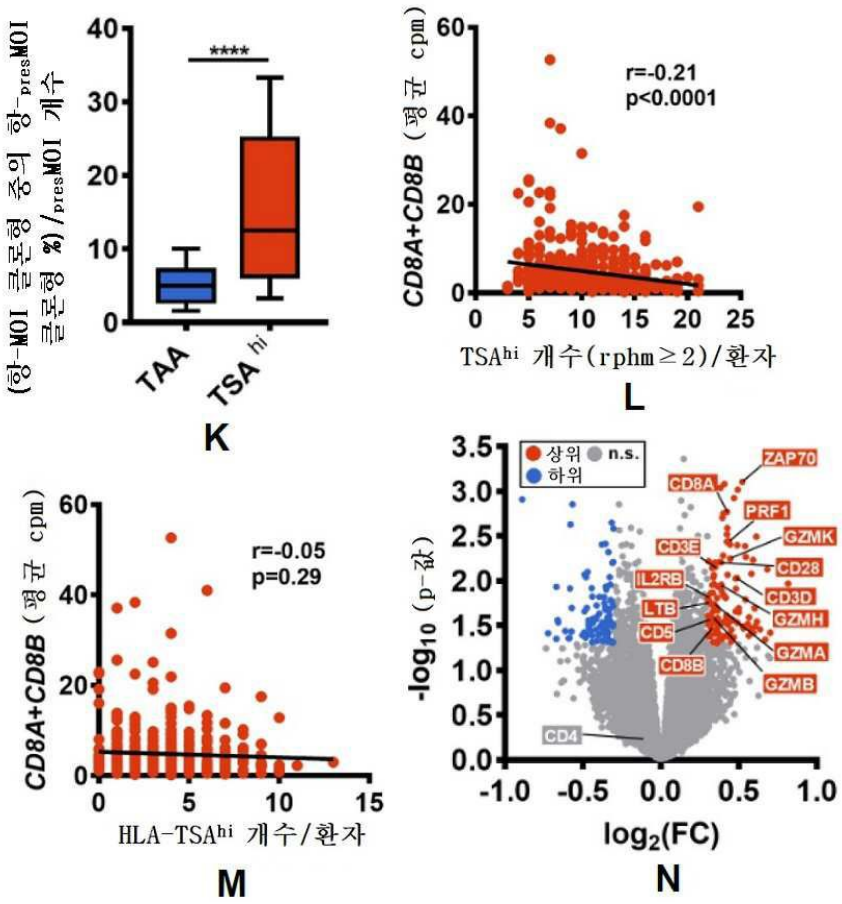

IFN-γ
TNF-α
IL-2
CD8+ T 세포의 %
무관한
SLLSGLLRA
펩타이드 없음
ALDPLLLRI
E


CD8
0,036
무관한
0,45
SLLSGLLRA
0,026
무관한
1,43
ALDPLLLRI
덱스트라머
F

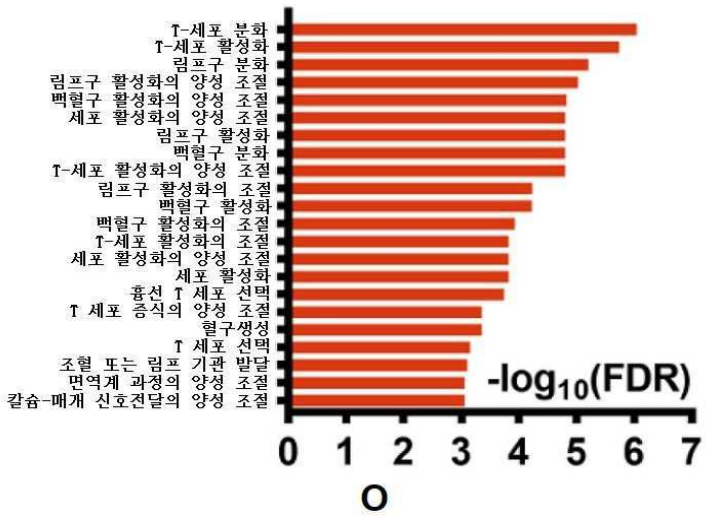
도면6gj


비배양됨
펩타이드 없음
TSA 풀
확장된 클론형의 풍부도
200
100
60
40
20
0
TSA 풀:
#1
#2
#3
G
TCR CPK
10
8
6
4
2
0
**
<15
≥15
<5
≥5
HLA-TAA 개수
HLA-TSA^hi 개수
H
모든 클론형 중의 항-TSA^hi 클론형 %
1.5
1.0
0.5
0.0
**
0-2
3-5
>5
HLA-TSA^hi 개수/환자
I
모든 클론형 중의 항-TAA 클론형 %
1.0
0.8
0.6
0.4
0.2
0.0
0-8
9-14
>15
HLA-TAA 개수/환자
J

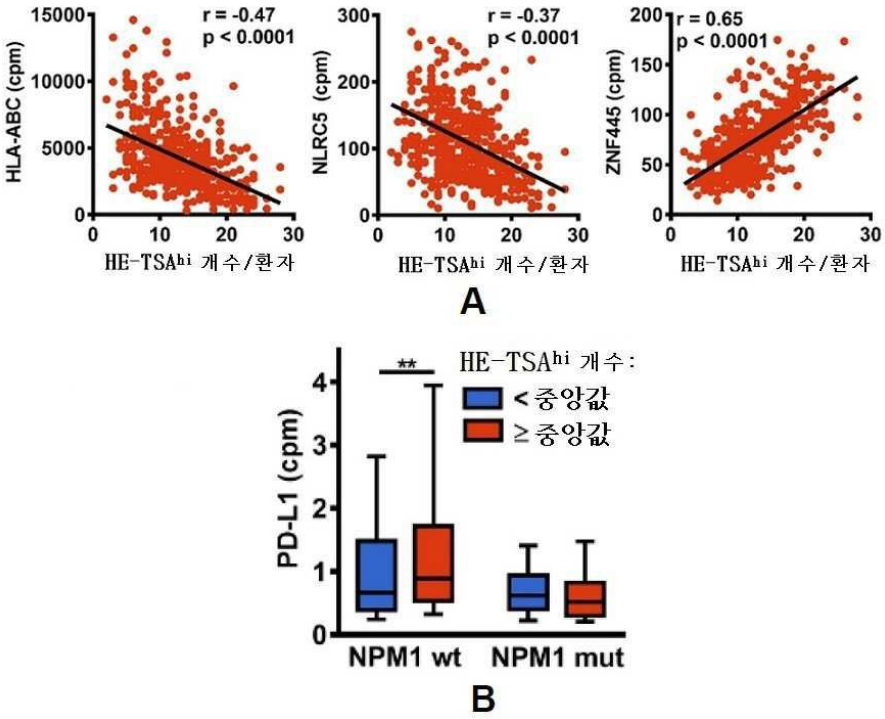
도면6kn


(항-MOI 클론형 중의 항-presMOI 클론형 %) /presMOI 개수
****
TAA
TSA hi
K
CD8A+CD8B (평균 cpm)
r=-0.21
p<0.0001
TSAhi 개수(rphm≥2)/환자
L
r=-0.05
p=0.29
HLA-TSAhi 개수/환자
M
상위
하위
n.s.
ZAP70
CD8A
PRF1
GZMK
CD3E
CD28
IL2RB
CD3D
LTB
GZMH
CD5
GZMA
CD8B
GZMB
CD4
-log10 (p-값)
log2(FC)
N

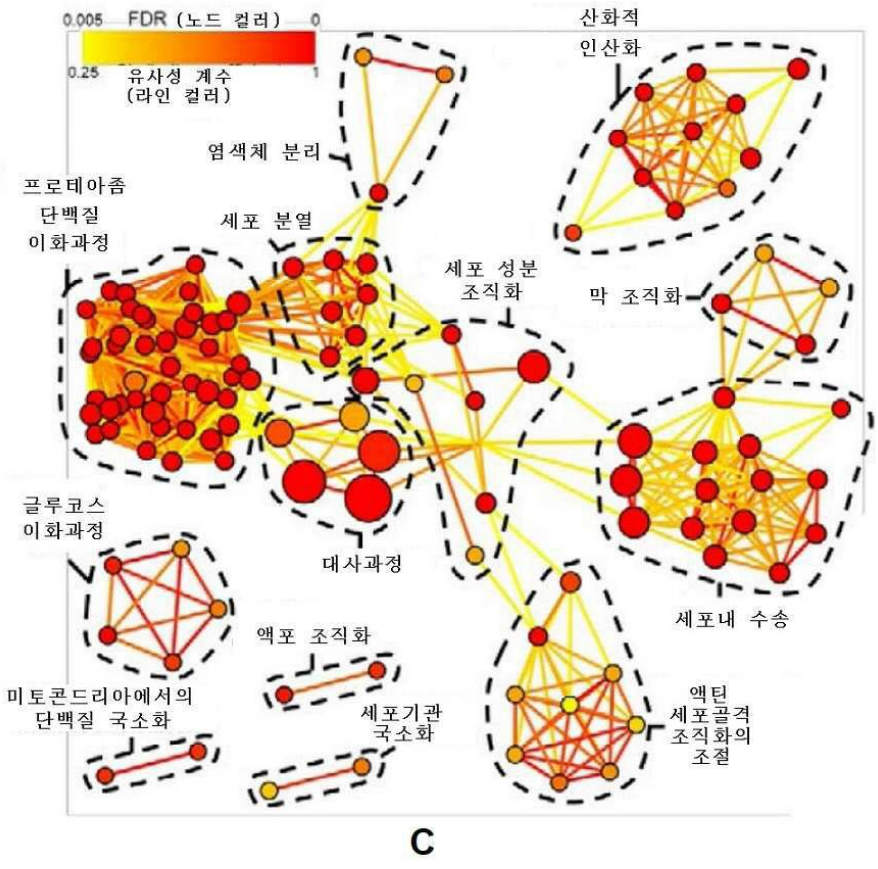
도면6o


T-세포 분화
T-세포 활성화
림프구 분화
림프구 활성화의 양성 조절
백혈구 활성화의 양성 조절
세포 활성화의 양성 조절
림프구 활성화
백혈구 분화
T-세포 활성화의 양성 조절
림프구 활성화의 조절
백혈구 활성화
백혈구 활성화의 조절
T-세포 활성화의 조절
세포 활성화의 양성 조절
세포 활성화
휴선 T 세포 선택
T 세포 증식의 양성 조절
혈구생성
T 세포 선택
조혈 또는 림프 기관 발달
면역계 과정의 양성 조절
칼슘-매개 신호전달의 양성 조절
-log10(FDR)
O

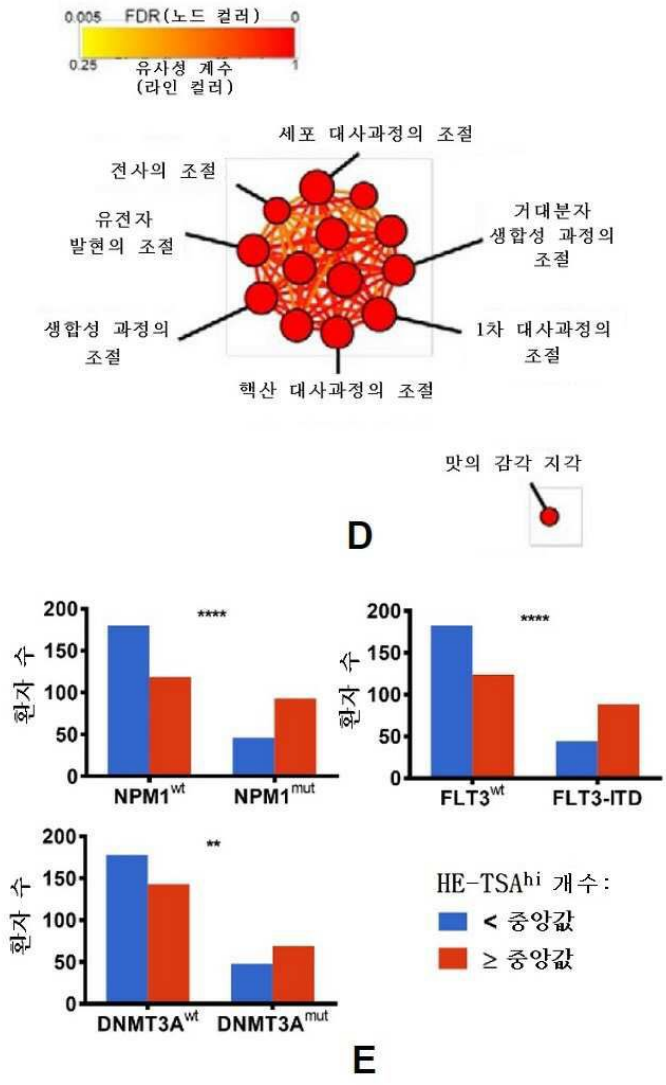
도면7ab


HLA-ABC (cpm)
r = -0.47
p < 0.0001
HE-TSAhi 개수/환자
NLRC5 (cpm)
r = -0.37
p < 0.0001
HE-TSAhi 개수/환자
ZNF445 (cpm)
r = 0.65
p < 0.0001
HE-TSAhi 개수/환자
A
PD-L1 (cpm)
**
HE-TSAhi 개수:
< 중앙값
≥ 중앙값
NPM1 wt
NPM1 mut
B

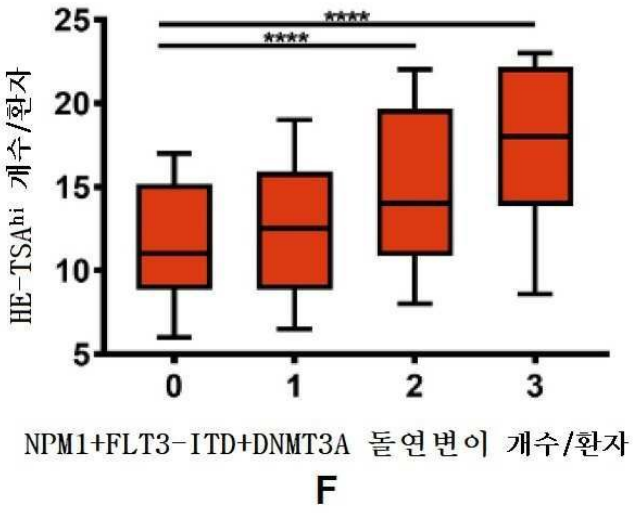
도면7c


FDR (노드 컬러)
0.005
0
0.25
1
유사성 계수
(라인 컬러)
염색체 분리
프로테아좀
단백질
이화과정
세포 분열
세포 성분
조직화
산화적
인산화
막 조직화
글루코스
이화과정
대사과정
세포내 수송
액포 조직화
미토콘드리아에서의
단백질 국소화
세포기관
국소화
액틴
세포골격
조직화의
조절
C

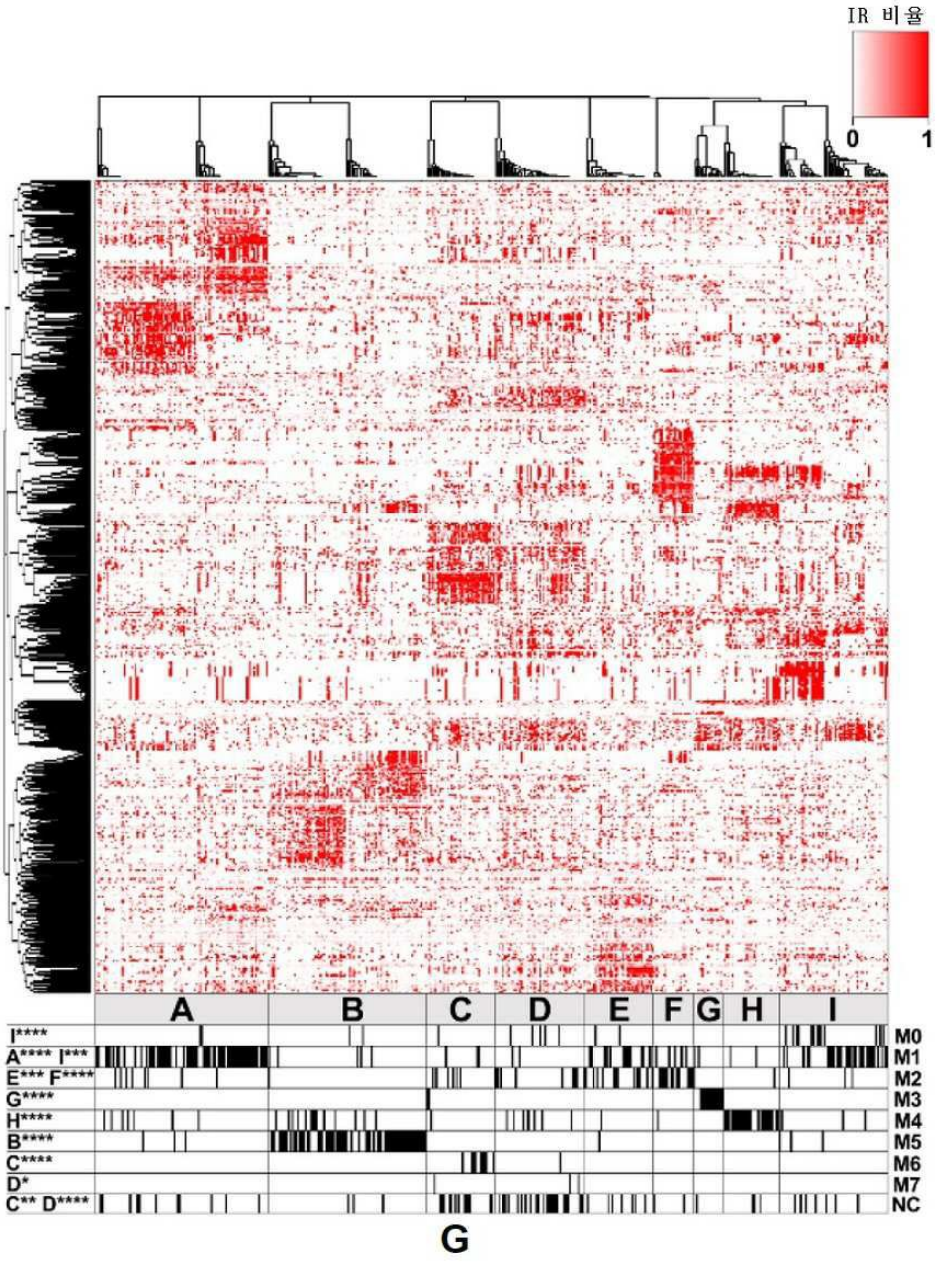
도면7de


0.005 FDR(노드 컬러) 0
0.25 유사성 계수 (라인 컬러) 1
세포 대사과정의 조절
전사의 조절
유전자 발현의 조절
생합성 과정의 조절
핵산 대사과정의 조절
1차 대사과정의 조절
거대분자 생합성 과정의 조절
맛의 감각 지각
D
환자 수
NPM1wt
NPM1mut
FLT3wt
FLT3-ITD
DNMT3Awt
DNMT3Amut
HE-TSAhi 개수:
< 중앙값
≥ 중앙값
E

도면7f

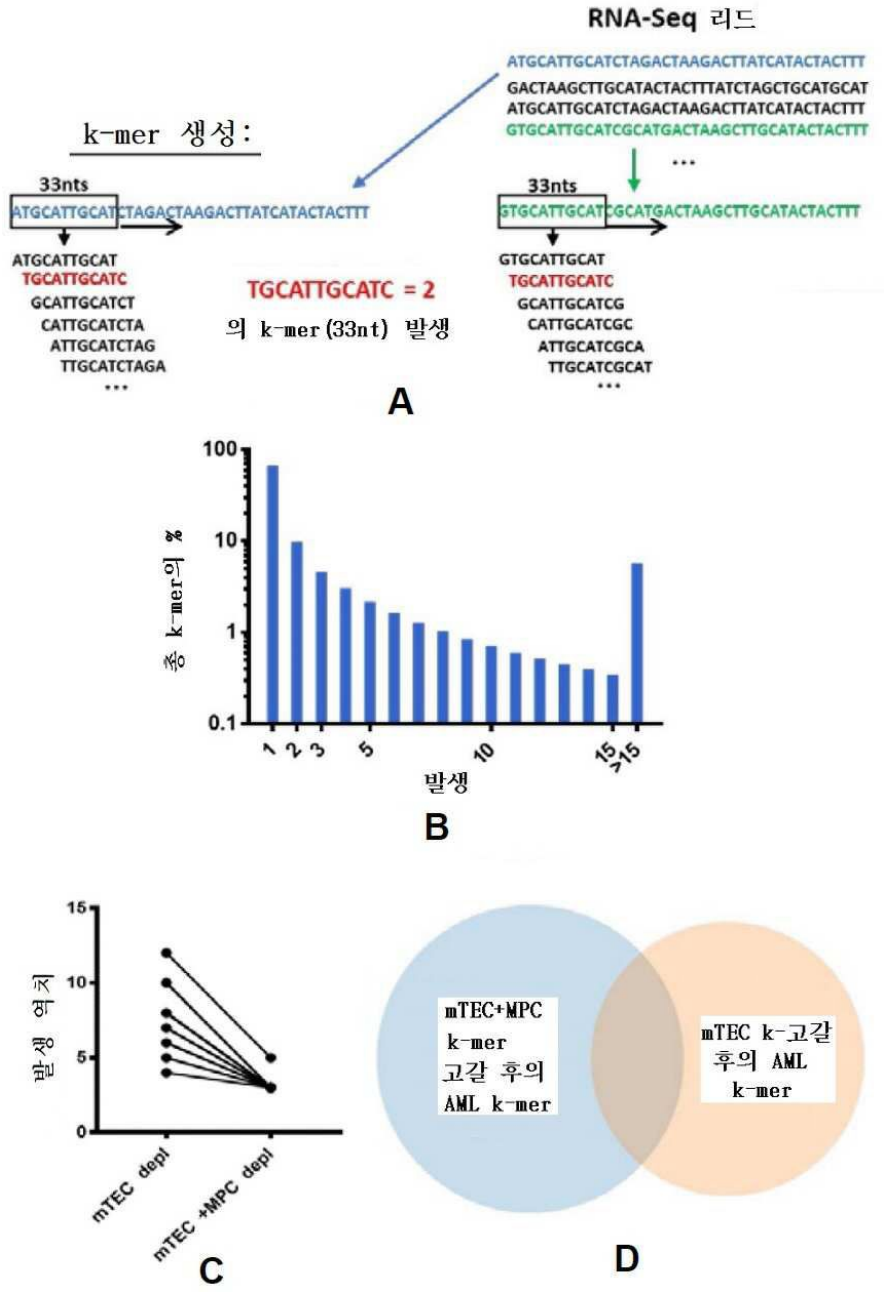

HE-TSAhi 개수/환자
NPM1+FLT3-ITD+DNMT3A 돌연변이 개수/환자
F

도면7g

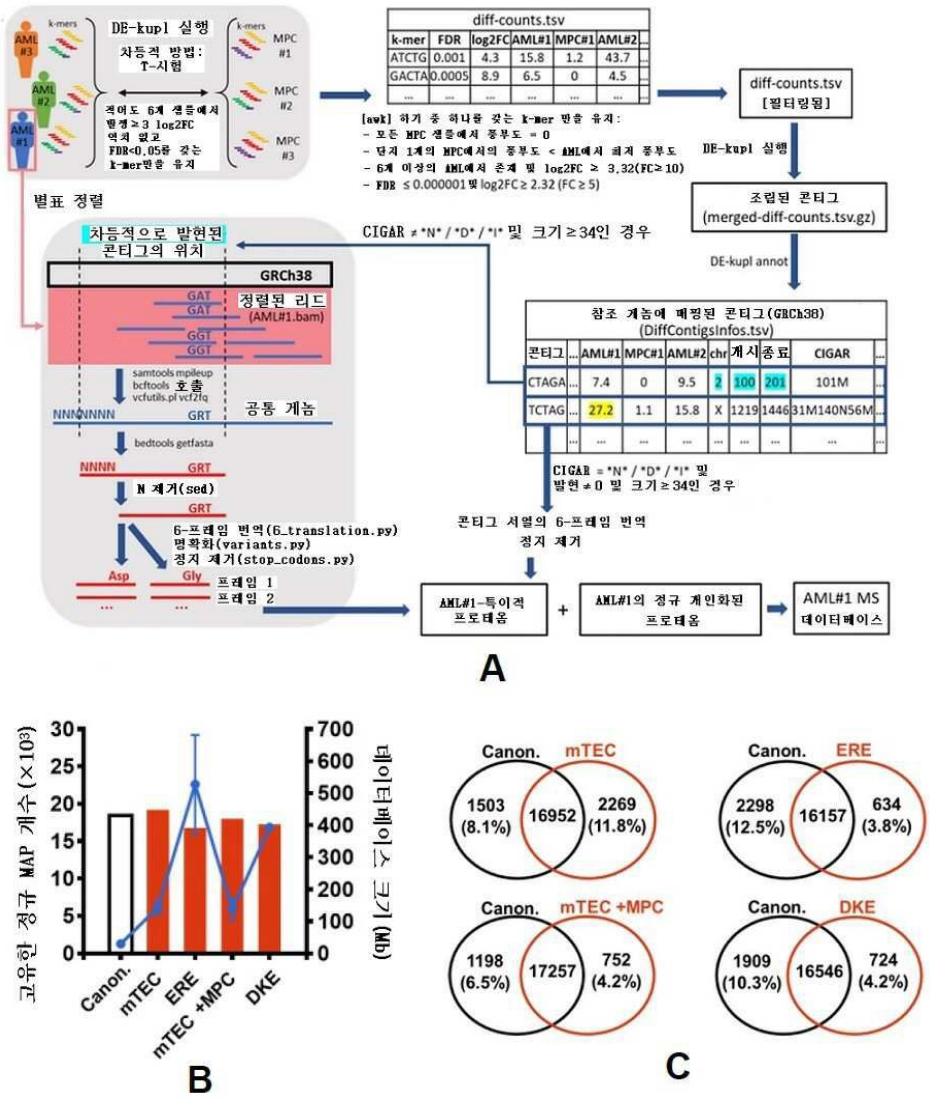


G

도면8

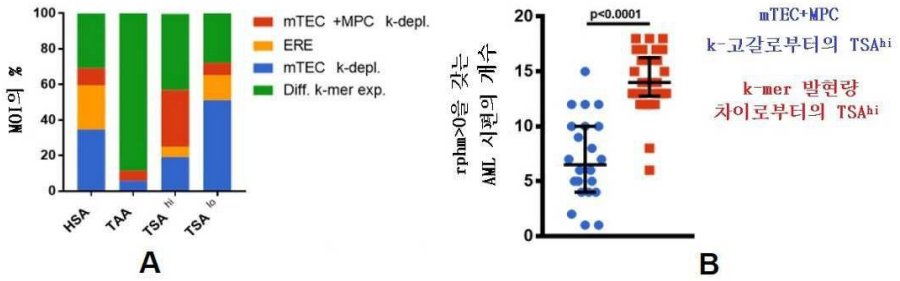

RNA-Seq 리드
ATGCATTGCATCTAGACTAAGACTTATCATACTACTTT
GACTAAGCTTGCATACTACTTTATCTAGCTGCATGCAT
ATGCATTGCATCTAGACTAAGACTTATCATACTACTTT
GTGCATTGCATCGCATGACTAAGCTTGCATACTACTTT
k-mer 생성:
33nts
ATGCATTGCATCTAGACTAAGACTTATCATACTACTTT
ATGCATTGCAT
TGCATTGCATC
GCATTGCATCT
CATTGCATCTA
ATTGCATCTAG
TTGCATCTAGA
TGCATTGCATC = 2
의 k-mer(33nt) 발생
33nts
GTGCATTGCATCGCATGACTAAGCTTGCATACTACTTT
GTGCATTGCAT
TGCATTGCATC
GCATTGCATCG
CATTGCATCGC
ATTGCATCGCA
TTGCATCGCAT
A
100
10
1
0.1
총 k-mer의 %
1
2
3
5
10
15
>15
발생
B
15
10
5
0
발생 역치
mTEC depl
mTEC +MPC depl
C
mTEC+MPC
k-mer
고갈 후의
AML k-mer
mTEC k-고갈
후의 AML
k-mer
D

도면9

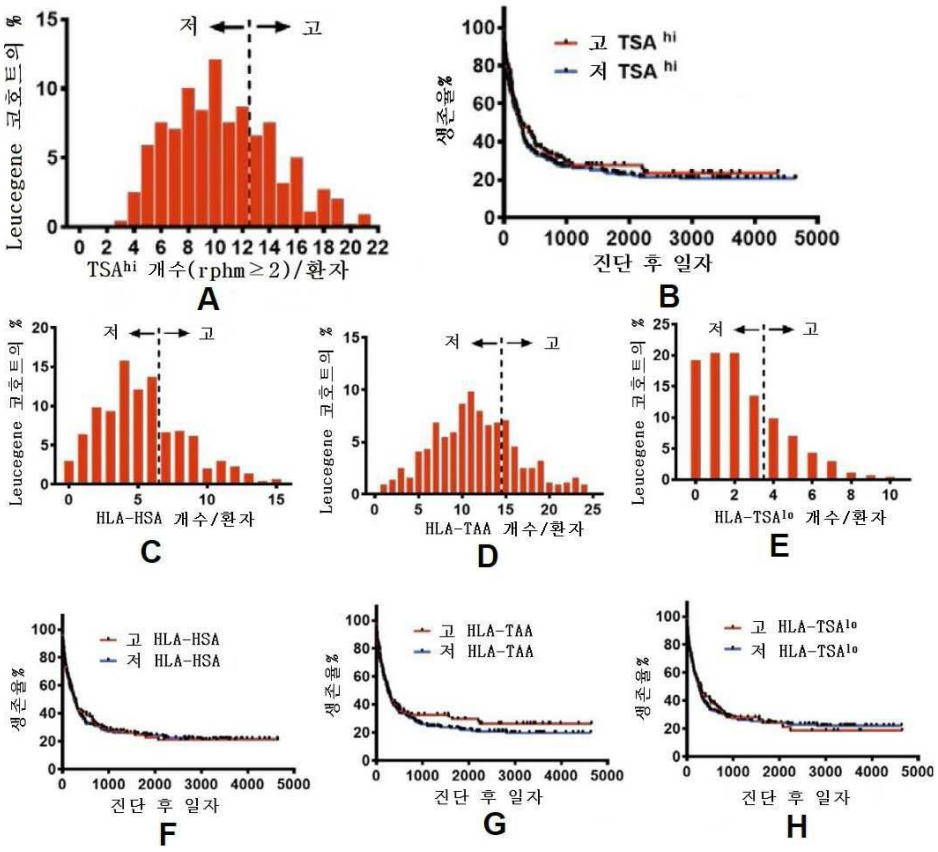

DE-kupl 실행
차등적 방법: T-시험
별표 정렬
diff-counts.tsv
diff-counts.tsv [필터링됨]
DE-kupl 실행
조립된 콘티그 (merged-diff-counts.tsv.gz)
DE-kupl annot
차등적으로 발현된 콘티그의 위치
GRCh38
정렬된 리드 (AML#1.bam)
공통 게놈
참조 게놈에 매핑된 콘티그(GRCh38) (DiffContigsInfos.tsv)
CIGAR ≠ *N* / *D* / *I* 및 크기 ≥ 34인 경우
CIGAR = *N* / *D* / *I* 및 발현 ≠ 0 및 크기 ≥ 34인 경우
콘티그 서열의 6-프레임 번역 정지 제거
N 제거(sed)
6-프레임 번역(6_translation.py)
명확화(variants.py)
정지 제거(stop_codons.py)
프레임 1
프레임 2
AML#1-특이적 프로테옴
AML#1의 정규 게인화된 프로테옴
AML#1 MS 데이터베이스
A

고유한 정규 MAP 개수 (×10³)
데이터베이스 크기 (Mb)
Canon.
mTEC
ERE
mTEC +MPC
DKE
B
1503 (8.1%)
16952
2269 (11.8%)
2298 (12.5%)
16157
634 (3.8%)
1198 (6.5%)
17257
752 (4.2%)
1909 (10.3%)
16546
724 (4.2%)
C

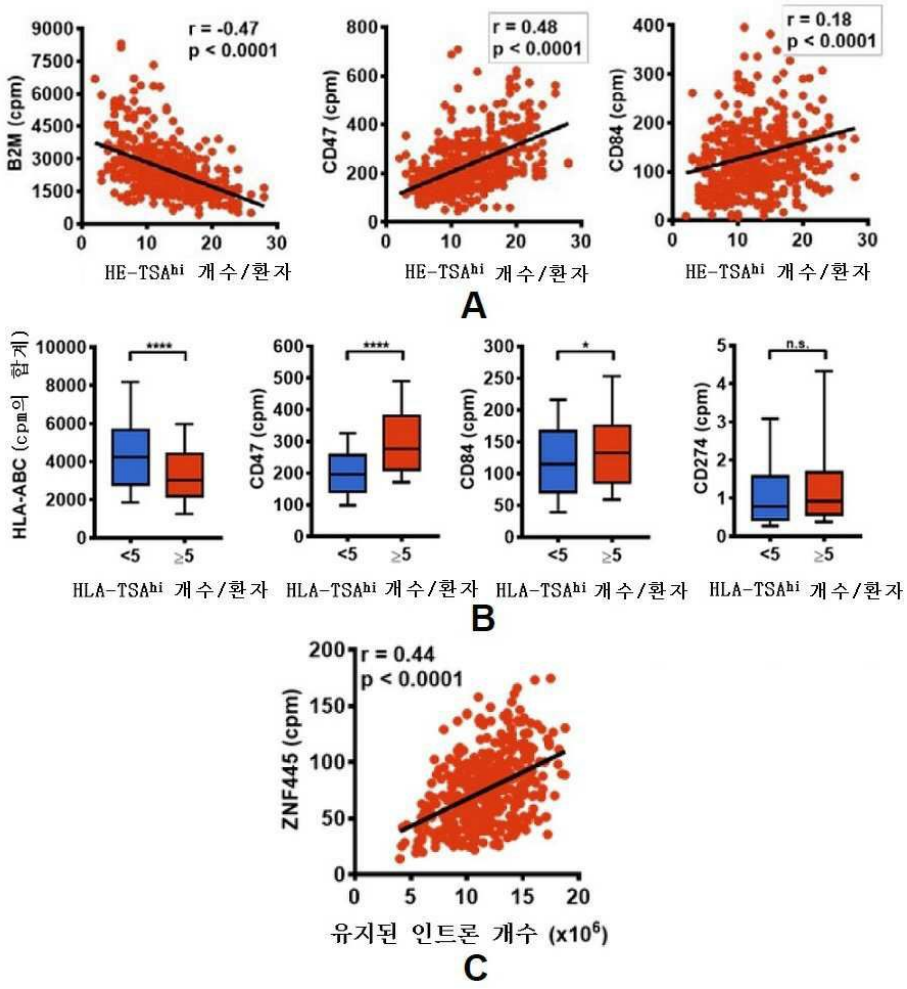
도면10


MOI의 %
mTEC +MPC k-depl.
ERE
mTEC k-depl.
Diff. k-mer exp.
HSA
TAA
TSA hi
TSA lo
A
rpkm>0을 갖는 AML 시편의 개수
p<0.0001
mTEC+MPC
k-고갈로부터의 TSAhi
k-mer 발현량 차이로부터의 TSAhi
B

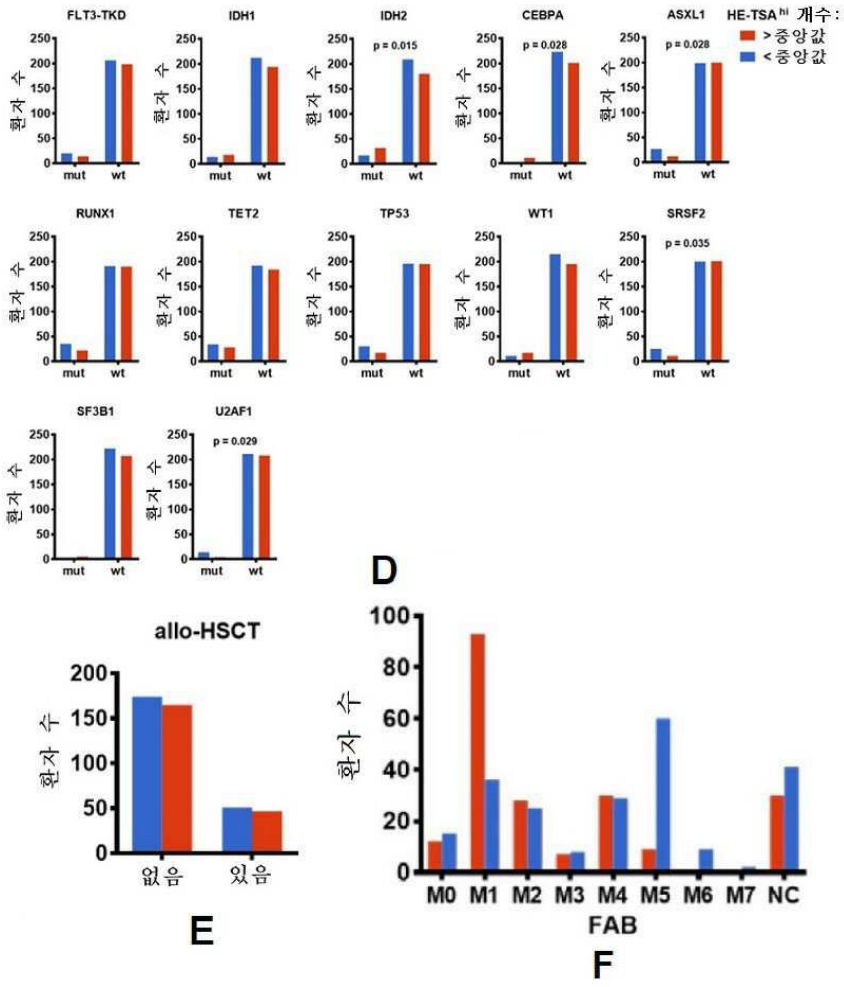
도면11


저
고
Leucegene 코호트의 %
TSA^hi 개수(rphm≥2)/환자
A
고 TSA hi
저 TSA hi
생존율%
진단 후 일자
B
HLA-HSA 개수/환자
C
HLA-TAA 개수/환자
D
HLA-TSA^lo 개수/환자
E
고 HLA-HSA
저 HLA-HSA
F
고 HLA-TAA
저 HLA-TAA
G
고 HLA-TSA^lo
저 HLA-TSA^lo
H

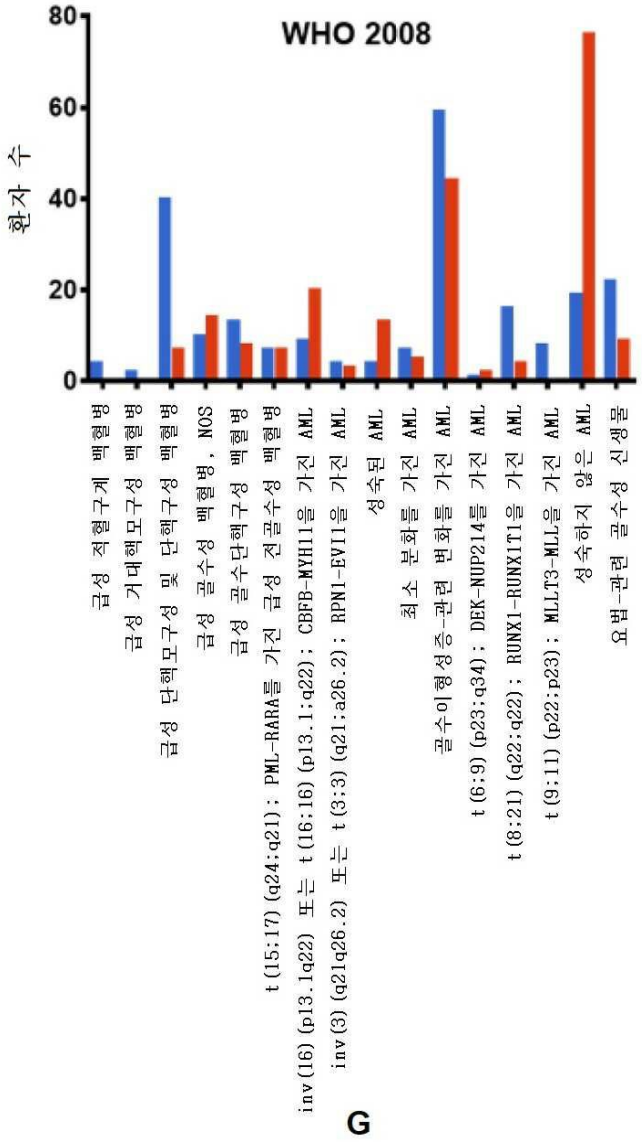
도면12ac


r = -0.47
p < 0.0001
B2M (cpm)
HE-TSAhi 개수/환자
r = 0.48
p < 0.0001
CD47 (cpm)
HE-TSAhi 개수/환자
r = 0.18
p < 0.0001
CD84 (cpm)
HE-TSAhi 개수/환자
A
HLA-ABC (cpm의 합계)
CD47 (cpm)
CD84 (cpm)
CD274 (cpm)
<5
≥5
HLA-TSAhi 개수/환자
B
r = 0.44
p < 0.0001
ZNF445 (cpm)
유지된 인트론 개수 (x10^6)
C

도면12df

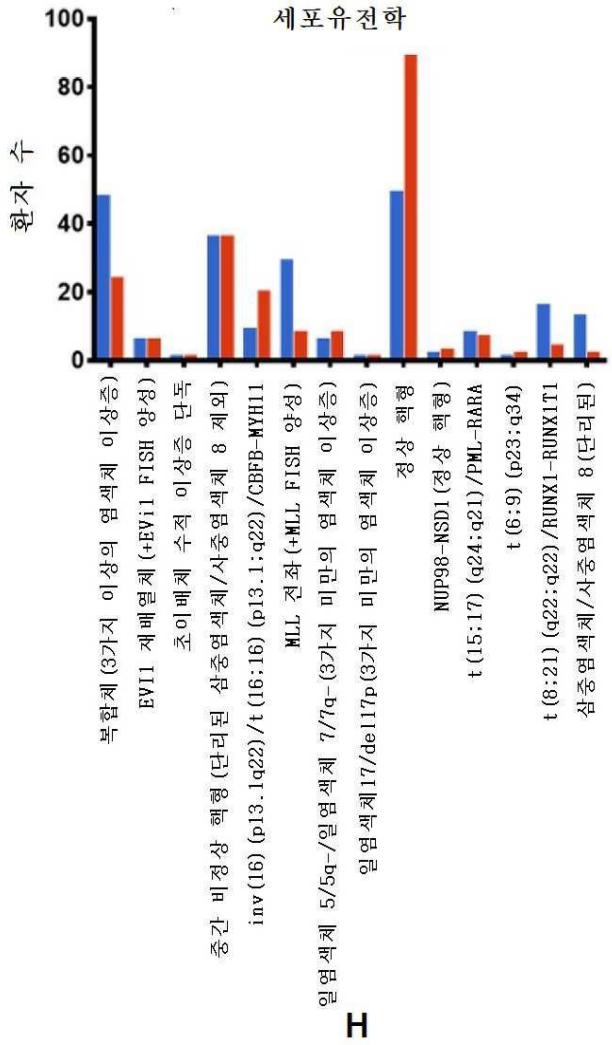

FLT3-TKD
IDH1
IDH2
p = 0.015
CEBPA
p = 0.028
ASXL1
p = 0.028
HE-TSA hi 개수:
> 중앙값
< 중앙값
RUNX1
TET2
TP53
WT1
SRSF2
p = 0.035
SF3B1
U2AF1
p = 0.029
환자 수
mut
wt
D
allo-HSCT
없음
있음
E
M0 M1 M2 M3 M4 M5 M6 M7 NC
FAB
F

도면12g


WHO 2008
환자 수
80
60
40
20
0
급성 적혈구계 백혈병
급성 거대핵모구성 백혈병
급성 단핵모구성 및 단핵구성 백혈병
급성 골수성 백혈병, NOS
급성 골수단핵구성 백혈병
t(15;17)(q24;q21); PML-RARA를 가진 급성 전골수성 백혈병
inv(16)(p13.1q22) 또는 t(16;16)(p13.1;q22); CBFB-MYH11을 가진 AML
inv(3)(q21q26.2) 또는 t(3;3)(q21;a26.2); RPN1-EVI1을 가진 AML
성숙된 AML
최소 분화를 가진 AML
골수이형성증-관련 변화를 가진 AML
t(6;9)(p23;q34); DEK-NUP214를 가진 AML
t(8;21)(q22;q22); RUNX1-RUNX1T1을 가진 AML
t(9;11)(p22;p23); MLLT3-MLL을 가진 AML
성숙하지 않은 AML
요법-관련 골수성 신생물

G

도면12h



서 열 목 록

SEQUENCE LISTING

<110> UNIVERSITE DE MONTREAL

<120> NOVEL TUMOR-SPECIFIC ANTIGENS FOR ACUTE MYELOID LEUKEMIA (AML) AND USES THEREOF

<130> G15691-00194

<140> PCT/CA2021/050340

<141> 2021-03-15

<150> US 63/009,853

<151> 2020-04-14

<160> 219

<170> PatentIn version 3.5

<210> 1

<211> 9

<212> PRT

<213> Homo sapiens

<400> 1

Arg Gln Ile Ser Val Gln Ala Ser Leu

1 5

<210> 2

<211> 9

<212> PRT

<213> Homo sapiens

<400> 2

Asp Arg Glu Leu Arg Asn Leu Glu Leu

1 5

<210> 3

<211> 9

<212> PRT

<213> Homo sapiens

<400> 3

Gly Ala Arg Gln Gln Ile His Ser Trp

1 5

<210> 4

<211> 8

<212> PRT

<213> Homo sapiens

<400> 4

Ser Gly Lys Leu Arg Val Ala Leu

1 5

<210> 5

<211> 11

<212> PRT

<213> Homo sapiens

<400> 5

Arg Ser Ala Ser Ser Ala Thr Gln Val His Lys

1 5 10

<210> 6

<211> 10

<212> PRT

<213> Homo sapiens

<400> 6

Ser Ala Ser Ser Ala Thr Gln Val His Lys

1 5 10

<210> 7

<211> 9

<212> PRT

<213> Homo sapiens

<400> 7

Phe Leu Leu Glu Phe Lys Pro Val Ser

1 5

<210> 8

<211> 8

<212> PRT

<213> Homo sapiens

<400> 8

Gly Pro Gln Val Arg Gly Ser Ile

1 5

<210> 9

<211> 8

<212> PRT

<213> Homo sapiens

<400> 9

Ile Arg Met Lys Ala Gln Ala Leu

1 5

<210> 10

<211> 9

<212> PRT

<213> Homo sapiens

<400> 10

Lys Ile Lys Val Phe Ser Lys Val Tyr

1 5

<210> 11

<211> 10

<212> PRT

<213> Homo sapiens

<400> 11

Leu Leu Ser Arg Gly Leu Leu Phe Arg Ile

1 5 10

<210> 12

<211> 9

<212> PRT

<213> Homo sapiens

<400> 12

Leu Pro Ile Ala Ser Ala Ser Leu Leu

1 5

<210> 13

<211> 9

<212>

PRT

<213> Homo sapiens

<400> 13

Leu Tyr Phe Leu Gly His Gly Ser Ile

1 5

<210> 14

<211> 9

<212> PRT

<213> Homo sapiens

<400> 14

Asn Pro Leu Gln Leu Ser Leu Ser Ile

1 5

<210> 15

<211> 8

<212> PRT

<213> Homo sapiens

<400> 15

Asp Leu Met Leu Arg Glu Ser Leu

1 5

<210> 16

<211> 8

<212> PRT

<213> Homo sapiens

<400> 16

Val Thr Phe Lys Leu Ser Leu Phe

1 5

<210> 17

<211> 8

<212> PRT

<213> Homo sapiens

<400> 17

Ile Ala Leu Tyr Lys Gln Val Leu

1 5

<210> 18

<211> 9

<212> PRT

<213> Homo sapiens

<400> 18

Ile Val Ala Thr Gly Ser Leu Leu Lys

1 5

<210> 19

<211> 9

<212> PRT

<213> Homo sapiens

<400> 19

Lys Ile Lys Asn Lys Thr Lys Asn Lys

1 5

<210> 20

<211> 9

<212> PRT

<213> Homo sapiens

<400> 20

Lys Leu Leu Ser Leu Thr Ile Tyr Lys

1 5

<210> 21

<211> 8

<212> PRT

<213> Homo sapiens

<400> 21

Asn Ile Leu Lys Lys Thr Val Leu

1 5

<210> 22

<211> 8

<212> PRT

<213> Homo sapiens

<400> 22

Asn Pro Lys Leu Lys Asp Ile Leu

1 5

<210> 23

<211> 8

<212> PRT

<213> Homo sapiens

<400> 23

Asn Gln Lys Lys Val Arg Ile Leu

1 5

<210> 24

<211> 10

<212> PRT

<213> Homo sapiens

<400> 24

Pro Phe Pro Leu Val Gln Val Glu Pro Val

1 5 10

<210> 25

<211> 8

<212> PRT

<213> Homo sapiens

<400> 25

Ser Pro Gln Ser Gly Pro Ala Leu

1 5

<210> 26

<211> 9

<212> PRT

<213> Homo sapiens

<400> 26

Thr Ser Arg Leu Pro Lys Ile Gln Lys

1 5

<210> 27

<211> 9

<212> PRT

<213> Homo sapiens

<400> 27

Leu Leu Asp Asn Ile Leu Gln Ser Ile

1 5

<210> 28

<211> 9

<212> PRT

<213> Homo sapiens

<400> 28

Arg Leu Glu Val Arg Lys Val Ile Leu

1 5

<210> 29

<211> 9

<212> PRT

<213> Homo sapiens

<400> 29

Leu Ser Trp Gly Tyr Phe Leu Phe Lys

1 5

<210> 30

<211> 9

<212> PRT

<213> Homo sapiens

<400> 30

Thr Ile Leu Pro Arg Ile Leu Thr Leu

1 5

<210> 31

<211> 8

<212> PRT

<213> Homo sapiens

<400> 31

Glu Gly Lys Ile Lys Arg Asn Ile

1 5

<210> 32

<211> 11

<212> PRT

<213> Homo sapiens

<400> 32

Phe Leu Ala Ser Phe Val Glu Lys Thr Val Leu

1 5 10

<210> 33

<211> 9

<212> PRT

<213> Homo sapiens

<400> 33

Ile Leu Ala Ser His Asn Leu Thr Val

1 5

<210> 34

<211> 9

<212> PRT

<213> Homo sapiens

<400> 34

Ile Gln Leu Thr Ser Val His Leu Leu

1 5

<210> 35

<211> 10

<212> PRT

<213> Homo sapiens

<400> 35

Leu Glu Leu Ile Ser Phe Leu Pro Val Leu

1 5 10

<210> 36

<211> 9

<212> PRT

<213> Homo sapiens

<400> 36

Asn Phe Cys Met Leu His Gln Ser Ile

1 5

<210> 37

<211> 8

<212> PRT

<213> Homo sapiens

<400> 37

Pro Ala Arg Pro Ala Gly Pro Leu

1 5

<210> 38

<211> 8

<212> PRT

<213> Homo sapiens

<400> 38

Pro Leu Pro Ile Val Pro Ala Leu

1 5

<210> 39

<211> 9

<212> PRT

<213> Homo sapiens

<400> 39

Ser Asn Leu Ile Arg Thr Gly Ser His

1 5

<210> 40

<211> 8

<212> PRT

<213> Homo sapiens

<400> 40

Val Pro Ala Pro Ala Gln Ala Ile

1 5

<210> 41

<211> 9

<212> PRT

<213> Homo sapiens

<400> 41

Lys Gly His Gly Gly Pro Arg Ser Trp

1 5

<210> 42

<211> 11

<212> PRT

<213> Homo sapiens

<400> 42

Ile Thr Ser Ser Ala Val Thr Thr Ala Leu Lys

1 5 10

<210> 43

<211> 9

<212> PRT

<213> Homo sapiens

<400> 43

Leu Leu Leu Pro Glu Ser Pro Ser Ile

1 5

<210> 44

<211> 9

<212> PRT

<213> Homo sapiens

<400> 44

Val Ile Leu Ile Pro Leu Pro Pro Lys

1 5

<210> 45

<211> 9

<212> PRT

<213> Homo sapiens

<400> 45

Ala Val Leu Leu Pro Lys Pro Pro Lys

1 5

<210> 46

<211> 9

<212> PRT

<213> Homo sapiens

<400> 46

Thr Gln Val Ser Met Ala Glu Ser Ile

1 5

<210> 47

<211> 8

<212> PRT

<213> Homo sapiens

<400> 47

Leu Asn His Leu Arg Thr Ser Ile

1 5

<210> 48

<211> 9

<212> PRT

<213> Homo sapiens

<400> 48

Asn Thr Ser His Leu Pro Leu Ile Tyr

1 5

<210> 49

<211> 8

<212> PRT

<213> Homo sapiens

<400> 49

Ser Ile Gln Arg Asn Leu Ser Leu

1 5

<210> 50

<211> 9

<212> PRT

<213> Homo sapiens

<400> 50

Asn Val Ser Ser His Val His Thr Val

1 5

<210> 51

<211> 9

<212> PRT

<213> Homo sapiens

<400> 51

Ala Leu Ala Ser His Leu Ile Glu Ala

1 5

<210> 52

<211> 9

<212> PRT

<213> Homo sapiens

<400> 52

Ala Leu Asp Asp Ile Thr Ile Gln Leu

1 5

<210> 53

<211> 9

<212> PRT

<

213> Homo sapiens

<400> 53

Ala Leu Gly Asn Thr Val Pro Ala Val

1 5

<210> 54

<211> 9

<212> PRT

<213> Homo sapiens

<400> 54

Ala Leu Leu Pro Ala Val Pro Ser Leu

1 5

<210> 55

<211> 9

<212> PRT

<213> Homo sapiens

<400> 55

Ala Pro Ala Pro Pro Pro Val Ala Val

1 5

<210> 56

<211> 8

<212> PRT

<213> Homo sapiens

<400> 56

Ala Pro Asp Lys Lys Ile Thr Leu

1 5

<210> 57

<211> 9

<212> PRT

<213> Homo sapiens

<400> 57

Ala Gln Met Asn Leu Leu Gln Lys Tyr

1 5

<210> 58

<211> 9

<212> PRT

<213> Homo sapiens

<400> 58

Asp Gln Val Ile Arg Leu Ala Gly Leu

1 5

<210> 59

<211> 9

<212> PRT

<213> Homo sapiens

<400> 59

Glu Thr Thr Ser Gln Val Arg Lys Tyr

1 5

<210> 60

<211> 9

<212> PRT

<213> Homo sapiens

<400> 60

Gly Gly Ser Leu Ile His Pro Gln Trp

1 5

<210> 61

<211> 9

<212> PRT

<213> Homo sapiens

<400> 61

Gly Leu Tyr Tyr Lys Leu His Asn Val

1 5

<210> 62

<211> 9

<212> PRT

<213> Homo sapiens

<400> 62

Gly Gln Lys Pro Val Ile Leu Thr Tyr

1 5

<210> 63

<211> 9

<212> PRT

<213> Homo sapiens

<400> 63

Gly Ser Leu Asp Phe Gln Arg Gly Trp

1 5

<210> 64

<211> 9

<212> PRT

<213> Homo sapiens

<400> 64

His His Leu Val Glu Thr Leu Lys Phe

1 5

<210> 65

<211> 9

<212> PRT

<213> Homo sapiens

<400> 65

His Leu Leu Ser Glu Thr Pro Gln Leu

1 5

<210> 66

<211> 9

<212> PRT

<213> Homo sapiens

<400> 66

His Gln Leu Tyr Arg Ala Ser Ala Leu

1 5

<210> 67

<211> 10

<212> PRT

<213> Homo sapiens

<400> 67

His Thr Asp Asp Ile Glu Asn Ala Lys Tyr

1 5 10

<210> 68

<211> 8

<212> PRT

<213> Homo sapiens

<400> 68

Ile Ala Ala Pro Ile Leu His Val

1 5

<210> 69

<211> 9

<212> PRT

<213> Homo sapiens

<400> 69

Lys Ala Phe Pro Phe His Ile Ile Phe

1 5

<210> 70

<211> 9

<212> PRT

<213> Homo sapiens

<400> 70

Lys Ala Thr Glu Tyr Val His Ser Leu

1 5

<210> 71

<211> 9

<212> PRT

<

213> Homo sapiens

<400> 71

Lys Phe Ser Asn Val Thr Met Leu Phe

1 5

<210> 72

<211> 9

<212> PRT

<213> Homo sapiens

<400> 72

Lys Leu Leu Glu Lys Ala Phe Ser Ile

1 5

<210> 73

<211> 9

<212> PRT

<213> Homo sapiens

<400> 73

Lys Pro Met Pro Thr Lys Val Val Phe

1 5

<210> 74

<211> 9

<212> PRT

<213> Homo sapiens

<400> 74

Asn Val Asn Arg Pro Leu Thr Met Lys

1 5

<210> 75

<211> 11

<212> PRT

<213> Homo sapiens

<400> 75

Arg Glu Pro Tyr Glu Leu Thr Val Pro Ala Leu

1 5 10

<210> 76

<211> 10

<212> PRT

<213> Homo sapiens

<400> 76

Ser Glu Ala Glu Ala Ala Lys Asn Ala Leu

1 5 10

<210> 77

<211> 9

<212> PRT

<213> Homo sapiens

<400> 77

Ser Leu Trp Gly Gln Pro Ala Glu Ala

1 5

<210> 78

<211> 11

<212> PRT

<213> Homo sapiens

<400> 78

Ser Pro Ala Asp His Arg Gly Tyr Ala Ser Leu

1 5 10

<210> 79

<211> 9

<212> PRT

<213> Homo sapiens

<400> 79

Ser Pro Gln Ser Ala Ala Ala Glu Leu

1 5

<210> 80

<211> 8

<212> PRT

<213> Homo sapiens

<400> 80

Ser Pro Val Val His Gln Ser Leu

1 5

<210> 81

<211> 8

<212> PRT

<213> Homo sapiens

<400> 81

Ser Pro Tyr Arg Thr Pro Val Leu

1 5

<210> 82

<211> 9

<212> PRT

<213> Homo sapiens

<400> 82

Ser Val Phe Ala Gly Val Val Gly Val

1 5

<210> 83

<211> 9

<212> PRT

<213> Homo sapiens

<400> 83

Ser Tyr Ser Pro Ala His Ala Arg Leu

1 5

<210> 84

<211> 9

<212> PRT

<213> Homo sapiens

<400> 84

Thr His Gly Ser Glu Gln Leu His Leu

1 5

<210> 85

<211> 9

<212> PRT

<213> Homo sapiens

<400> 85

Thr Gln Ala Pro Pro Asn Val Val Leu

1 5

<210> 86

<211> 10

<212> PRT

<213> Homo sapiens

<400> 86

Val Leu Val Pro Tyr Glu Pro Pro Gln Val

1 5 10

<210> 87

<211> 9

<212> PRT

<213> Homo sapiens

<400> 87

Val Ser Phe Pro Asp Val Arg Lys Val

1 5

<210> 88

<211> 11

<212> PRT

<213> Homo sapiens

<400> 88

Val Val Phe Asp Lys Ser Asp Leu Ala Lys Tyr

1 5 10

<210> 89

<211> 9

<212> PRT

<213> Homo sapiens

<400> 89

Tyr Ser His His Ser Gly Leu Glu Tyr

1 5

<210> 90

<211> 9

<212> PRT

<213> Homo sapiens

<400> 90

Tyr Tyr Leu Asp Trp Ile His His Tyr

1 5

<210> 91

<211> 9

<212> PRT

<213> Homo sapiens

<400> 91

Ser Val Tyr Lys Tyr Leu Lys Ala Lys

1 5

<210> 92

<211> 9

<212> PRT

<213> Homo sapiens

<400> 92

Ile Tyr Gln Phe Ile Met Asp Arg Phe

1 5

<210> 93

<211> 9

<212> PRT

<213> Homo sapiens

<400> 93

Gly Thr Leu Gln Gly Ile Arg Ala Trp

1 5

<210> 94

<211> 10

<212> PRT

<213> Homo sapiens

<400> 94

Ala Gln Lys Val Ser Val Gly Gln Ala Ala

1 5 10

<210> 95

<211> 9

<212> PRT

<213> Homo sapiens

<400> 95

Leu Tyr Pro Ser Lys Leu Thr His Phe

1 5

<210> 96

<211> 9

<

212> PRT

<213> Homo sapiens

<400> 96

Ala Thr Gln Asn Thr Ile Ile Gly Lys

1 5

<210> 97

<211> 9

<212> PRT

<213> Homo sapiens

<400> 97

Ala Gln Asp Ile Ile Leu Gln Ala Val

1 5

<210> 98

<211> 9

<212> PRT

<213> Homo sapiens

<400> 98

Pro Pro Arg Pro Leu Gly Ala Gln Val

1 5

<210> 99

<211> 9

<212> PRT

<213> Homo sapiens

<400> 99

Phe Asn Val Ala Leu Asn Ala Arg Tyr

1 5

<210> 100

<211> 9

<212> PRT

<213> Homo sapiens

<400> 100

Gly Pro Gly Ser Arg Glu Ser Thr Leu

1 5

<210> 101

<211> 8

<212> PRT

<213> Homo sapiens

<400> 101

Ile Pro His Gln Arg Ser Ser Leu

1 5

<210> 102

<211> 9

<212> PRT

<213> Homo sapiens

<400> 102

Leu Thr Asp Arg Ile Tyr Leu Thr Leu

1 5

<210> 103

<211> 9

<212> PRT

<213> Homo sapiens

<400> 103

Asn Leu Lys Glu Lys Lys Ala Leu Phe

1 5

<210> 104

<211> 10

<212> PRT

<213> Homo sapiens

<400> 104

Val Leu Phe Gly Gly Lys Val Ser Gly Ala

1 5 10

<210> 105

<211> 9

<212> PRT

<213> Homo sapiens

<400> 105

Val Val Phe Pro Phe Pro Val Asn Lys

1 5

<210> 106

<211> 9

<212> PRT

<213> Homo sapiens

<400> 106

Ser Leu Leu Ile Ile Pro Lys Lys Lys

1 5

<210> 107

<211> 9

<212> PRT

<213> Homo sapiens

<400> 107

Ala Pro Gly Ala Ala Gly Gln Arg Leu

1 5

<210> 108

<211> 9

<212> PRT

<213> Homo sapiens

<400> 108

Lys Leu Gln Asp Lys Glu Ile Gly Leu

1 5

<210> 109

<211> 9

<212> PRT

<213> Homo sapiens

<400> 109

Ser Leu Arg Glu Pro Gln Pro Ala Leu

1 5

<210> 110

<211> 9

<212

> PRT

<213> Homo sapiens

<400> 110

Thr Pro Gly Arg Ser Thr Gln Ala Ile

1 5

<210> 111

<211> 8

<212> PRT

<213> Homo sapiens

<400> 111

Ala Pro Arg Gly Thr Ala Ala Leu

1 5

<210> 112

<211> 8

<212> PRT

<213> Homo sapiens

<400> 112

Ile Ala Ser Pro Ile Ala Leu Leu

1 5

<210> 113

<211> 9

<212> PRT

<213> Homo sapiens

<400> 113

Ile Leu Phe Gln Asn Ser Ala Leu Lys

1 5

<210> 114

<211> 8

<212> PRT

<213> Homo sapiens

<400> 114

Ile Leu Lys Lys Asn Ile Ser Ile

1 5

<210> 115

<211> 8

<212> PRT

<213> Homo sapiens

<400> 115

Ile Pro Leu Ala Val Arg Thr Ile

1 5

<210> 116

<211> 8

<212> PRT

<213> Homo sapiens

<400> 116

Leu Pro Arg Asn Lys Pro Leu Leu

1 5

<210> 117

<211> 9

<212> PRT

<213> Homo sapiens

<400> 117

Pro Ala Pro Pro His Pro Ala Ala Leu

1 5

<210> 118

<211> 8

<212> PRT

<213> Homo sapiens

<400> 118

Ser Pro Val Val Arg Val Gly Leu

1 5

<210> 119

<211> 10

<212> PRT

<213> Homo sapiens

<400> 119

Thr Leu Asn Gln Gly Ile Asn Val Tyr Ile

1 5 10

<210> 120

<211> 9

<212> PRT

<213> Homo sapiens

<400> 120

Arg Pro Arg Gly Pro Arg Thr Ala Pro

1 5

<210> 121

<211> 11

<212> PRT

<213> Homo sapiens

<400> 121

Ser Val Gln Leu Leu Glu Gln Ala Ile His Lys

1 5 10

<210> 122

<211> 9

<212> PRT

<213> Homo sapiens

<400> 122

Arg Thr Pro Lys Asn Tyr Gln His Trp

1 5

<210> 123

<211> 9

<212> PRT

<213> Homo sapiens

<400> 123

Ala Leu Pro Val Ala Leu Pro Ser Leu

1 5

<210> 124

<211> 9

<212> PRT

<213> Homo sapiens

<400> 124

Ser Leu Gln Ile Leu Val Ser Ser Leu

1 5

<210> 125

<211> 9

<212> PRT

<213> Homo sapiens

<400> 125

Ile Ser Asn Lys Val Pro Lys Leu Phe

1 5

<210> 126

<211> 10

<212> PRT

<213> Homo sapiens

<400> 126

Thr Val Ile Arg Ile Ala Ile Val Asn Lys

1 5 10

<210> 127

<211> 9

<212> PRT

<213> Homo sapiens

<400> 127

Lys Glu Ile Phe Leu Glu Leu Arg Leu

1 5

<210> 128

<211> 9

<212> PRT

<213> Homo sapiens

<400> 128

Thr Leu Arg Ser Pro Gly Ser Ser Leu

1 5

<210> 129

<211> 9

<212> PRT

<213> Homo sapiens

<400> 129

Thr Val Arg Gly Asp Val Ser Ser Leu

1 5

<210> 130

<211> 9

<212> PRT

<213> Homo sapiens

<400> 130

Ala Leu Asp Pro Leu Leu Leu Arg Ile

1 5

<210> 131

<211> 9

<212> PRT

<213> Homo sapiens

<400> 131

Ile Ser Leu Ile Val Thr Gly Leu Lys

1 5

<210> 132

<211> 9

<212> PRT

<213> Homo sapiens

<400> 132

Lys Ile Leu Asp Val Asn Leu Arg Ile

1 5

<210> 133

<211> 9

<212> PRT

<213> Homo sapiens

<400> 133

Glu Arg Val Tyr Ile Arg Ala Ser Leu

1 5

<210> 134

<211> 8

<212> PRT

<213> Homo sapiens

<400> 134

Ile Leu Asp Leu Glu Ser Arg Tyr

1 5

<210> 135

<211> 9

<212> PRT

<213> Homo sapiens

<400> 135

Lys Thr Phe Val Gln Gln Lys Thr Leu

1 5

<210> 136

<211> 9

<212> PRT

<213> Homo sapiens

<400> 136

Leu Tyr Ile Lys Ser Leu Pro Ala Leu

1 5

<210> 137

<211> 9

<212> PRT

<213> Homo sapiens

<400> 137

Val Leu Lys Glu Lys Asn Ala Ser Leu

1 5

<210> 138

<211> 9

<212> PRT

<213> Homo sapiens

<400> 138

Leu Gly Ile Ser Leu Thr Leu Lys Tyr

1 5

<210> 139

<211> 8

<212> PRT

<213> Homo sapiens

<400> 139

Asp Leu Leu Pro Lys Lys Leu Leu

1 5

<210> 140

<211> 9

<212> PRT

<213> Homo sapiens

<400> 140

His Ser Leu Ile Ser Ile Val Tyr Leu

1 5

<210> 141

<211> 9

<212> PRT

<213> Homo sapiens

<400> 141

Ile Ala Gly Ala Leu Arg Ser Val Leu

1 5

<210> 142

<211> 9

<212> PRT

<213> Homo sapiens

<400> 142

Ile Gly Asn Pro Ile Leu Arg Val Leu

1 5

<210> 143

<211> 9

<212> PRT

<213> Homo sapiens

<400> 143

Ile Tyr Ala Pro His Ile Arg Leu Ser

1 5

<210> 144

<211> 8

<212> PRT

<213> Homo sapiens

<400> 144

Leu Arg Ser Gln Ile Leu Ser Tyr

1 5

<210> 145

<211> 9

<212> PRT

<213> Homo sapiens

<400> 145

Arg Tyr Leu Ala Asn Lys Ile His Ile

1 5

<210> 146

<211> 9

<212> PRT

<213> Homo sapiens

<400> 146

Ser Leu Leu Ser Gly Leu Leu Arg Ala

1 5

<210> 147

<211> 8

<212> PRT

<213> Homo sapiens

<400> 147

Ser Arg Ile His Leu Val Val Leu

1 5

<210> 148

<211> 10

<212> PRT

<213> Homo sapiens

<400> 148

Ser Ser Ser Pro Val Arg Gly Pro Ser Val

1 5 10

<210> 149

<211> 9

<212

> PRT

<213> Homo sapiens

<400> 149

Ser Thr Phe Ser Leu Tyr Leu Lys Lys

1 5

<210> 150

<211> 9

<212> PRT

<213> Homo sapiens

<400> 150

Ser Leu Asp Leu Leu Pro Leu Ser Ile

1 5

<210> 151

<211> 9

<212> PRT

<213> Homo sapiens

<400> 151

Val Thr Asp Leu Leu Ala Leu Thr Val

1 5

<210> 152

<211> 11

<212> PRT

<213> Homo sapiens

<400> 152

Arg Thr Gln Ile Thr Lys Val Ser Leu Lys Lys

1 5 10

<210> 153

<211> 8

<212> PRT

<213> Homo sapiens

<400> 153

Ile Leu Arg Ser Pro Leu Lys Trp

1 5

<210> 154

<211> 9

<212> PRT

<213> Homo sapiens

<400> 154

Leu Ser Thr Gly His Leu Ser Thr Val

1 5

<210> 155

<211> 9

<212> PRT

<213> Homo sapiens

<400> 155

Thr Val Glu Glu Tyr Leu Val Asn Ile

1 5

<210> 156

<211> 10

<212

> PRT

<213> Homo sapiens

<400> 156

Gln Ile Lys Thr Lys Leu Leu Gly Ser Leu

1 5 10

<210> 157

<211> 10

<212> PRT

<213> Homo sapiens

<400> 157

Leu Pro Ser Phe Ser His Phe Leu Leu Leu

1 5 10

<210> 158

<211> 9

<212> PRT

<213> Homo sapiens

<400> 158

Cys Leu Arg Ile Gly Pro Val Thr Leu

1 5

<210> 159

<211> 9

<212> PRT

<213> Homo sapiens

<400> 159

His Val Ser Asp Gly Ser Thr Ala Leu

1 5

<210> 160

<211> 9

<212> PRT

<213> Homo sapiens

<400> 160

Ile Ala Tyr Ser Val Arg Ala Leu Arg

1 5

<210> 161

<211> 8

<212> PRT

<213> Homo sapiens

<400> 161

Pro Arg Gly Phe Leu Ser Ala Leu

1 5

<210> 162

<211> 9

<212> PRT

<213> Homo sapiens

<400> 162

Ile Ser Ser Trp Leu Ile Ser Ser Leu

1 5

<210> 163

<211> 9

<212> PRT

<213> Homo sapiens

<400> 163

Ile Pro Leu Asn Pro Phe Ser Ser Leu

1 5

<210> 164

<211> 8

<212> PRT

<213> Homo sapiens

<400> 164

Leu Ser Asp Arg Gln Leu Ser Leu

1 5

<210> 165

<211> 9

<212> PRT

<213> Homo sapiens

<400> 165

Leu Ser His Pro Ala Pro Ser Ser Leu

1 5

<210> 166

<211> 9

<212> PRT

<213> Homo sapiens

<400> 166

Leu Arg Lys Ala Val Asp Pro Ile Leu

1 5

<210> 167

<211> 9

<212> PRT

<213> Homo sapiens

<400> 167

Ile Leu Leu Glu Glu Gln Ser Leu Ile

1 5

<210> 168

<211> 9

<212> PRT

<213> Homo sapiens

<400> 168

Leu Thr Ser Ile Ser Ile Arg Pro Val

1 5

<210> 169

<211> 9

<212> PRT

<213> Homo sapiens

<400> 169

Thr Ile Ser Glu Cys Pro Leu Leu Ile

1 5

<210> 170

<211> 9

<212> PRT

<213> Homo sapiens

<400> 170

Thr Leu Lys Leu Lys Lys Ile Phe Phe

1 5

<210> 171

<211> 9

<212> PRT

<213> Homo sapiens

<400> 171

Ile Leu Leu Ser Asn Phe Ser Ser Leu

1 5

<210> 172

<211> 9

<212> PRT

<213> Homo sapiens

<400> 172

Leu Gly Gly Ala Trp Lys Ala Val Phe

1 5

<210> 173

<211> 9

<212> PRT

<213> Homo sapiens

<400> 173

Leu Ser Ala Ser His Leu Ser Ser Leu

1 5

<210> 174

<211> 9

<212> PRT

<213> Homo sapiens

<400> 174

Ala Gly Asp Ile Ile Ala Arg Leu Ile

1 5

<210> 175

<211> 9

<212> PRT

<213> Homo sapiens

<400> 175

Asp Arg Gly Ile Leu Arg Asn Leu Leu

1 5

<210> 176

<211> 9

<212> PRT

<213> Homo sapiens

<400> 176

Gly Leu Arg Leu Ile His Val Ser Leu

1 5

<210> 177

<211> 9

<212> PRT

<213> Homo sapiens

<400> 177

Gly Leu Arg Leu Leu His Val Ser Leu

1 5

<210> 178

<211> 9

<212> PRT

<213> Homo sapiens

<400> 178

Leu His Asn Glu Lys Gly Leu Ser Leu

1 5

<210> 179

<211> 11

<212> PRT

<213> Homo sapiens

<400> 179

Leu Pro Ser Phe Ser Arg Pro Ser Gly Ile Ile

1 5 10

<210> 180

<211> 9

<212> PRT

<213> Homo sapiens

<400> 180

Leu Ser Ser Arg Leu Pro Leu Gly Lys

1 5

<210> 181

<211> 8

<212> PRT

<213> Homo sapiens

<400> 181

Met Ile Gly Ile Lys Arg Leu Leu

1 5

<210> 182

<211> 8

<212> PRT

<213> Homo sapiens

<400> 182

Asn Leu Lys Lys Arg Glu Ile Leu

1 5

<210> 183

<211> 9

<212> PRT

<213> Homo sapiens

<400> 183

Arg Met Val Ala Tyr Leu Gln Gln Leu

1 5

<210> 184

<211> 9

<212> PRT

<213> Homo sapiens

<400> 184

Ser Pro Ala Arg Ala Leu Pro Ser Leu

1 5

<210> 185

<211> 8

<212> PRT

<213> Homo sapiens

<400> 185

Thr Val Pro Gly Ile Gln Arg Tyr

1 5

<210> 186

<211> 9

<212> PRT

<213> Homo sapiens

<400> 186

Val Ser Arg Asn Tyr Val Leu Leu Ile

1 5

<210> 187

<211> 9

<212> PRT

<213> Homo sapiens

<400> 187

Leu Thr Val Pro Leu Ser Val Phe Trp

1 5

<210> 188

<211> 9

<212> PRT

<213> Homo sapiens

<400> 188

Lys Leu Asn Gln Ala Phe Leu Val Leu

1 5

<210> 189

<211> 10

<212> PRT

<213> Homo sapiens

<400> 189

Arg Leu Val Ser Ser Thr Leu Leu Gln Lys

1 5 10

<210> 190

<211> 8

<212> PRT

<213> Homo sapiens

<400> 190

Leu Pro Ser His Ser Leu Leu Ile

1 5

<210> 191

<211> 13

<212> PRT

<213> Homo sapiens

<400> 191

Cys Ser Ala Arg Gly Asp Arg Glu Tyr Glu Gln Tyr Phe

1 5 10

<210> 192

<211> 15

<212> PRT

<213> Homo sapiens

<400> 192

Cys Ala Ser Thr Val Val Ala Gly Asn Ser Ser Pro Leu His Phe

1 5 10 15

<210> 193

<211> 16

<212> PRT

<213> Homo sapiens

<400> 193

Cys Ala Ser Ser Gln Asp Gly Ile Trp Gly Ala Tyr Glu Gln Tyr Phe

1 5 10 15

<210> 194

<211> 15

<212> PRT

<213> Homo sapiens

<400> 194

Cys Ala Ser Ser Val Asp Ala Gly Gly Asn Tyr Glu Gln Tyr Phe

1 5 10 15

<210> 195

<211> 16

<212> PRT

<213> Homo sapiens

<400> 195

Cys Ala Ser Ser Leu Gly Gly Gln Gly Leu Ser Tyr Gly Tyr Thr Phe

1 5 10 15

<210> 196

<211> 12

<212> PRT

<213> Homo sapiens

<400> 196

Cys Ala Ser Ser Tyr Arg Pro Asn Glu Gln Tyr Phe

1 5 10

<210> 197

<211> 14

<212> PRT

<213> Homo sapiens

<400> 197

Cys Ala Ser Ser Gly Thr Asp Leu Asn Gln Pro Gln His Phe

1 5 10

<210> 198

<211> 12

<212> PRT

<213> Homo sapiens

<400> 198

Cys Ala Ser Ser Ser Asn Phe Glu Pro Leu His Phe

1 5 10

<210> 199

<211> 15

<212> PRT

<213> Homo sapiens

<400> 199

Cys Ala Ser Ser Trp Gly Gly Ser Asn Thr Gly Glu Leu Phe Phe

1 5 10 15

<210> 200

<211> 15

<212> PRT

<213> Homo sapiens

<400> 200

Cys Ala Ser Ser Glu Tyr Arg Ala Leu Asn Thr Glu Ala Phe Phe

1 5 10 15

<210> 201

<211> 15

<212> PRT

<213> Homo sapiens

<400> 201

Cys Ala Ser Ser Glu Ser Gly Thr Gly Gly Gln Pro Gln His Phe

1 5 10 15

<210> 202

<211> 14

<212> PRT

<213> Homo sapiens

<400> 202

Cys Ala Ser Ser Arg Thr Gly Glu Asn Thr Glu Ala Phe Phe

1 5 10

<210> 203

<211> 15

<212> PRT

<213> Homo sapiens

<400> 203

Cys Ala Ser Ser Ser Thr Asp Arg Gln His Tyr Gly Tyr Thr Phe

1 5 10 15

<210> 204

<211> 17

<212> PRT

<213> Homo sapiens

<400> 204

Cys Ala Ser Ser Asp Arg Thr Gly Gly Ser Ser Asn Glu Lys Leu Phe

1 5 10 15

Phe

<210> 205

<211> 15

<212> PRT

<213> Homo sapiens

<400> 205

Cys Ala Thr Ser Arg Ser Gly Asp Ser Asn Gln Pro Gln His Phe

1 5 10 15

<210> 206

<211> 13

<212> PRT

<213> Homo sapiens

<400> 206

Cys Ala Ser Ser Tyr Val Leu Asn Thr Glu Ala Phe Phe

1 5 10

<210> 207

<211> 15

<212> PRT

<213> Homo sapiens

<400> 207

Cys Ala Ser Arg Glu Ser Gly Gln Met Asn Glu Lys Leu Phe Phe

1 5 10 15

<210> 208

<211> 15

<212> PRT

<213> Homo sapiens

<400> 208

Cys Ala Ser Ser Asp Gly Gly Glu Gly Thr Tyr Gly Tyr Thr Phe

1 5 10 15

<210> 209

<211> 17

<212> PRT

<213> Homo sapiens

<400> 209

Cys Ala Ser Thr Ser Trp Thr Gly Phe Gly Pro Asn Tyr Gly Tyr Thr

1 5 10 15

Phe

<210> 210

<211> 14

<212> PRT

<213> Homo sapiens

<400> 210

Cys Ala Ser Ser Ala Asp Ser Ser Leu Gly Gly Tyr Thr Phe

1 5 10

<210> 211

<211> 15

<212> PRT

<213> Homo sapiens

<400> 211

Cys Ser Ala Arg Asp Leu Ala Gly Gly Thr Tyr Glu Gln Tyr Phe

1 5 10 15

<210> 212

<211> 15

<212> PRT

<213> Homo sapiens

<400> 212

Cys Ala Ser Ser Pro Pro Thr Gly Glu Tyr Glu Lys Leu Phe Phe

1 5 10 15

<210> 213

<211> 9

<212> PRT

<213> Homo sapiens

<400> 213

Cys Ala Arg Ser Phe Gly Gly Phe Phe

1 5

<210> 214

<211> 15

<212> PRT

<213> Homo sapiens

<400> 214

Cys Ala Ser Ser Leu Pro Ser Gly Ile Leu Tyr Glu Gln Tyr Phe

1 5 10 15

<210> 215

<211> 14

<212> PRT

<213> Homo sapiens

<400> 215

Cys Ala Ser Ala Pro Gly Gly Met Pro Tyr Gly Tyr Thr Phe

1 5 10

<210> 216

<211> 15

<212> PRT

<213> Homo sapiens

<400> 216

Cys Ala Ser Ser Leu Gln Asp Thr Asp Tyr Asn Glu Gln Phe Phe

1 5 10 15

<210> 217

<211> 18

<212> PRT

<213> Homo sapiens

<400> 217

Cys Ala Ser Ser Asp Pro Gly Thr Ser Gly Val Phe Thr Gly Glu Leu

1 5 10 15

Phe Phe

<210> 218

<211> 12

<212> PRT

<213> Homo sapiens

<400> 218

Cys Ser Ala Arg Leu Gly Thr Gly Glu Leu Phe Phe

1 5 10

<210> 219

<211> 12

<212> PRT

<213> Homo sapiens

<400> 219

Cys Ser Ala Arg Leu Gly Thr Gly Glu Leu Phe Phe

1 5 10